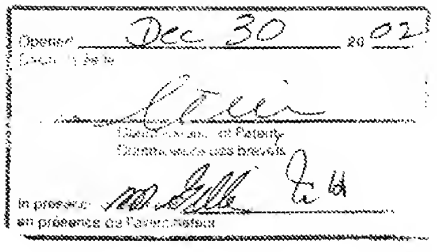


This is EXHIBIT FIERS-37
to
the Affidavit of Walter C. Fiers
sworn before me
this 19 th day of November, 2001

Commissioner for Oath or Notary Public

Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone

David V. Goeddel*, Herbert L. Heyneker*, Toyohara Hozumi†, Rene Arentzen†, Keiichi Itakura†, Daniel G. Yansura*, Michael J. Ross*, Giuseppe Miozzari*, Roberto Crea‡ & Peter H. Seeburg*

Divisions of Molecular Biology* and Organic Chemistry‡, Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080
† Division of Biology, City of Hope National Medical Center, Duarte, California 91010

DNA coding for human growth hormone was constructed by using chemically synthesised DNA in conjunction with enzymatically prepared cDNA. This 'hybrid' gene was expressed in Escherichia coli *under the control of the* lac *promoter. A polypeptide was produced having the size and immunological properties characteristic of mature human growth hormone.*

HUMAN GROWTH HORMONE (HGH) is a protein of 191 amino acids which is synthesised in the anterior lobe of the pituitary. Growth in hypopituitary dwarfs, whose small stature is due to a deficiency of HGH, can be restored during childhood by administration of this hormone[1]. In addition, HGH may prove effective in the treatment of a variety of ailments, including bone fractures, skin burns and bleeding ulcers[2]. As growth hormone is species specific, human cadavers have been the only source of HGH.

The primary translation product of growth hormone mRNA is a precursor protein consisting of a signal peptide attached to the N-terminus of growth hormone[3]. Such signal or 'pre' sequences are characteristic of secreted proteins. In the case of rat growth hormone (RGH), DNA sequencing of cDNA prepared from RGH mRNA has identified the 26 amino acid residues of the pre-sequence[4]. From information derived from the cDNA sequence of human growth hormone[5] we have designed and constructed a bacterial plasmid which instructs the synthesis of substantial quantities of mature HGH in a microbial cell.

HGH gene assembly strategy

The general approach used for the bacterial synthesis of HGH involves a combination of the cloning of the complementary DNA (cDNA) prepared from pituitary mRNA with the cloning of chemically synthesised DNA. The specific strategy was based on the known restriction endonuclease pattern of HGH cDNA[4]. *Hae*III sites are present in the 3' noncoding region and in the sequence coding for amino acids 23 and 24 of HGH. Treatment of double stranded (ds) HGH cDNA with *Hae*III gives a DNA fragment of 551 base pairs which includes coding sequences for amino acids 24–191 of HGH. We planned to clone this cDNA fragment and to clone separately a chemically synthesised DNA 'adaptor' fragment containing an ATG initiation codon and coding sequences for residues 1–23 of HGH. These two DNA fragments would be combined to form a synthetic–natural 'hybrid' gene. When inserted into a plasmid downstream from a suitable bacterial promoter and ribosome binding site, this gene could be expected to direct the synthesis of fMet-HGH. The fact that most bacterial proteins do not contain N-terminal methionine residues suggests that the fMet should be efficiently removed, resulting in the direct expression of HGH.

There are several advantages of this approach over conventional expression methods which utilise either chemically synthesised DNA or cDNA exclusively. The cDNA approach has been used to express several fused proteins (β-lactamase–rat proinsulin[6], β-lactamase–rat pre-growth hormone[7] and β-galactosidase–chicken ovalbumin[8,9]) in *Escherichia coli* which still retain antigenic activity, but which cannot be easily processed to give the natural gene product. However, the cloning of mouse dihydrofolate reductase cDNA into G-tailed, *Prd*-cleaved pBR322 resulted in the expression of a protein which appeared to be dihydrofolate reductase[10]. This protein could have been derived from the proteolytic cleavage of a β-lactamase–dihydrofolate reductase fused protein or from the direct initiation of translation at the dihydrofolate reductase ATG (AUG) start codon. Neither of these possibilities could be relied on for HGH expression. *In vivo* proteolytic cleavage of a fused bacterially produced protein to give HGH is highly unlikely, and initiation of translation at an HGH start codon would be expected to yield the pre-hormone containing an extra 26 amino acids. The alternative approach of complete chemical synthesis of genes, although shown to be extremely effective for the bacterial production of the human peptides somatostatin[11] and insulin[12], would be very time consuming for a polypeptide as large as HGH.

Construction and cloning of plasmids containing HGH DNA sequences

The plan for the chemical synthesis of DNA designed to code for the first 24 amino acids of HGH is shown in Fig. 1*a*. To facilitate joining to HGH cDNA, the synthetic plan incorporates the same *Hae*III restriction site at amino acid residues 23 and 24 that is found in the cDNA. The 12 indicated deoxyoligonucleotides (Fig. 1*a*) were synthesised by the improved phosphotriester method essentially as described previously for insulin[12]. Figure 1*b* summarises the steps involved in the assembly and cloning of these DNA fragments in pBR322 (ref. 15). Many transformants of *E. coli* 294 (ref. 16) were obtained with plasmids containing *Eco*RI–*Hin*dIII inserts of approximately the desired size. The nucleotide sequences of three of these inserts were determined. One, pHGH3, had the expected DNA sequence. The other two had a deletion of the A·T base pair at the third nucleotide position for the codon of amino acid 5 (Pro).

The construction of the plasmid (pHGH31) containing the *Hae*III cDNA fragment of HGH is outlined in Fig. 2. Total poly(A)-mRNA from human pituitaries was used in the synthesis of ds-cDNA by RNA-dependent DNA polymerase (reverse transcriptase) and DNA polymerase I Klenow fragment. The cDNA was treated with *Hae*III restriction endonuclease and DNA of about 550 base pairs was purified by gel electrophoresis. This DNA fraction was then 'tailed' with deoxyC residues using terminal deoxynucleotidyl transferase[13] and cloned in pBR322 DNA which had been cleaved with *Prd* and extended with deoxyG residues. This procedure restores the *Hae*III restriction sites on the cDNA and regenerates *Prd* sites at both ends of the insert. Approximately 200 transformants of *E. coli* χ1776 (ref. 20) were obtained from 1 ng of C-tailed cDNA. These colonies were screened for the presence of HGH

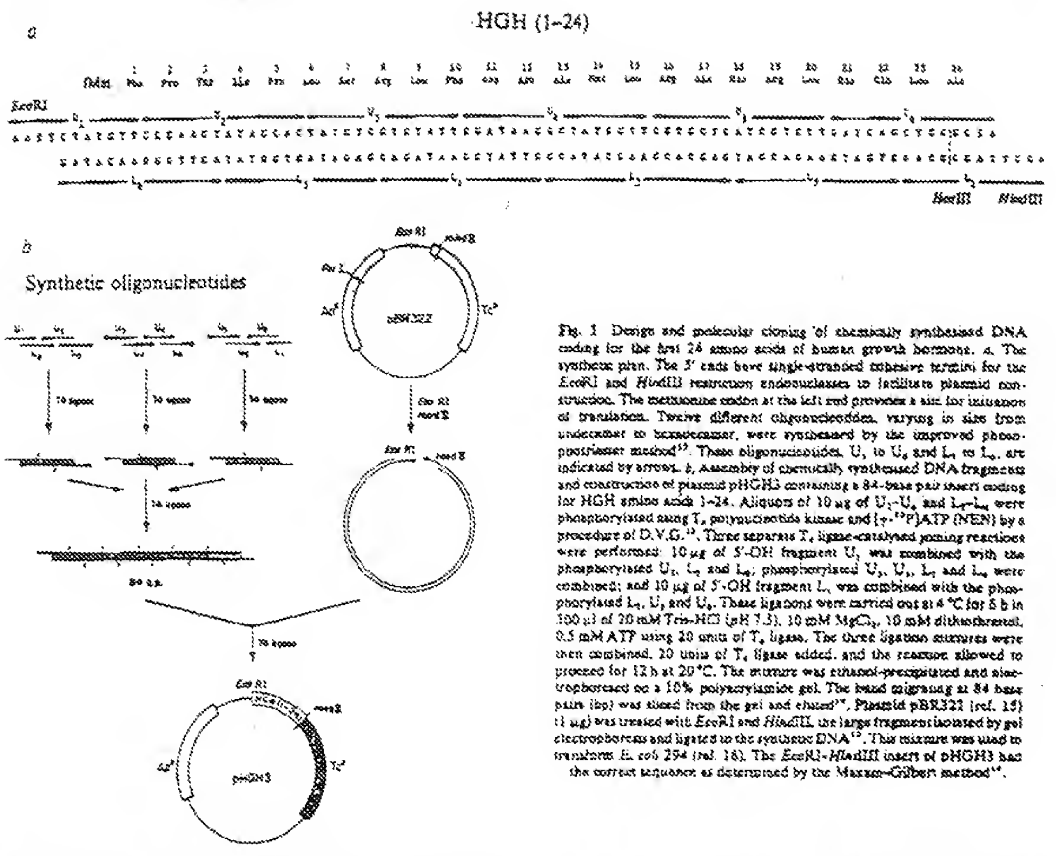


Fig. 1 Design and molecular cloning of chemically synthesized DNA coding for the first 24 amino acids of human growth hormone. *a*, The synthetic plan. The 5′ ends have single-stranded cohesive termini for the *Eco*RI and *Hin*dIII restriction endonucleases to facilitate plasmid construction. The methionine codon at the left end provides a site for initiation of translation. Twelve different oligonucleotides, U_1 to U_6 and L_1 to L_6, are indicated by arrows. *b*, Assembly of chemically synthesized DNA fragments and construction of plasmid pHGH3 containing a 84-base pair insert coding for HGH amino acids 1–24. Aliquots of 10 μg of U_1–U_6 and L_1–L_6 were phosphorylated using T_4 polynucleotide kinase and [γ-^{32}P]ATP (NEN) by a procedure of D.V.G.[19]. Three separate T_4 ligase-catalysed joining reactions were performed: 10 μg of 5′-OH fragment U_1 was combined with the phosphorylated U_2, U_3 and L_3; phosphorylated U_4, U_5, L_5 and L_6 were combined; and 10 μg of 5′-OH fragment L_4 was combined with the phosphorylated L_2, U_3 and U_6. These ligations were carried out at 4 °C for 5 h in 100 μl of 20 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.5 mM ATP using 20 units of T_4 ligase. The three ligation mixtures were then combined, 20 units of T_4 ligase added, and the reaction allowed to proceed for 12 h at 20 °C. The mixture was ethanol-precipitated and electrophoresed on a 10% polyacrylamide gel. The band migrating at 84 base pairs (bp) was eluted from the gel and eluted[20]. Plasmid pBR322 (ref. 15) (1 μg) was treated with *Eco*RI and *Hin*dIII, the large fragment isolated by gel electrophoresis and ligated to the synthetic DNA[15]. This mixture was used to transform *E. coli* 294 (ref. 16). The *Eco*RI–*Hin*dIII insert of pHGH3 had the correct sequence as determined by the Maxam–Gilbert method[14].

sequences by the Grunstein–Hogness hybridisation procedure[21] using ^{32}P-labelled human chorionic somatomammotropin cDNA[22] as probe. Seven colonies hybridized specifically with the probe, which is nearly identical to HGH in nucleotide sequence[3]. Three had identical *Hae*III inserts of approximately 550 base pairs containing *Sma*I, *Bgl*II, *Pst*I and *Pvu*II sites, all characteristic of HGH cDNA[3]. Sequence analysis of the insert from one of these clones, pHGH31 (Fig. 3) gave only one discrepancy (CUG versus CUA codon for leucine 101) when compared with the previously determined cDNA sequence[3]. The amino acid sequence derived from the cDNA sequence differs from the published amino acid sequences of LI[23] at positions 74 (Glu versus Gln), 107 (Asp versus Asn) and 109 (Asn versus Asp). Although we cannot rule out post-transcriptional modification, the discrepancies probably stem from shortcomings of earlier protein sequencing methods in the identification of amidated amino acids.

Construction and cloning of HGH expression plasmids

To express directly HGH in *E. coli*, a plasmid (pGH6) having two *lac* promoters was constructed as follows. A 285-base pair *Eco*RI fragment containing two 95-base pair UV-5 *lac* promoter fragments separated by a 95-base pair heterologous DNA fragment was isolated from the plasmid pKB268 (ref. 24) and inserted into pBR322 at the *Eco*RI site. A plasmid (pGH1) was isolated with the promoter orientated to initiate transcription in the direction of the tetracycline resistance (tc^R) gene. The *Eco*RI site distal to the tc^R gene was destroyed by a previously outlined procedure[11].

Figure 4 reviews the steps involved in assembling the two separately cloned HGH DNA fragments. Treatment of pHGH3 with *Eco*RI and *Hae*III gives a 77-base pair fragment with an *Eco*RI cohesive end and a *Hae*III blunt end (also see Fig. 1*a*). This fragment, which codes for HGH amino acids 1–23, was isolated from a polyacrylamide gel. The plasmid pHGH31 was cleaved with *Hae*III and the 55-base pair HGH cDNA fragment isolated by gel electrophoresis. *Xma*I treatment of this fragment yielded a 512-base pair fragment coding for amino acids 24–191 of HGH with one *Hae*III blunt end and one *Xma*I sticky end. The ligation of these two fragments resulted in the formation of high molecular weight DNA. Treatment with *Sma*I (which recognizes the same sequence as *Xma*I but leave flush ends) and *Eco*RI converted the ligation products to three distinct fragments: a 154-base pair dimer of the 77-base pair DNA fragment, a 1,024-base pair dimer of the 512-base pair fragment, and the desired 591-base pair DNA coding for the entire HGH sequence. This latter fragment was purified by gel electrophoresis and inserted into pGH6 between the *Eco*RI and the S_1 nuclease-treated *Hin*dIII sites. Following transformation of χ1776, the expression plasmid pHGH107 was identified by colony screening and restriction analysis as described in Fig. 3. The HGH DNA sequence of pHGH107 was verified by the Maxam–Gilbert procedure[14].

In our plan to express HGH two base pairs were included between the *Eco*RI AATT 'sticky' end and the ATG codon of the chemically synthesized portion of the HGH coding sequence (see Fig. 1). Therefore, in the plasmid pHGH107, 11 base pairs separate the *lac* AGGA ribosome binding site[26] and the ATG translational start for HGH:

5′...AGGAAACAGAATTCTATG...

3′...TCCTTTGTCTTAAGATAC...

In the naturally occurring *lac* system the separation between

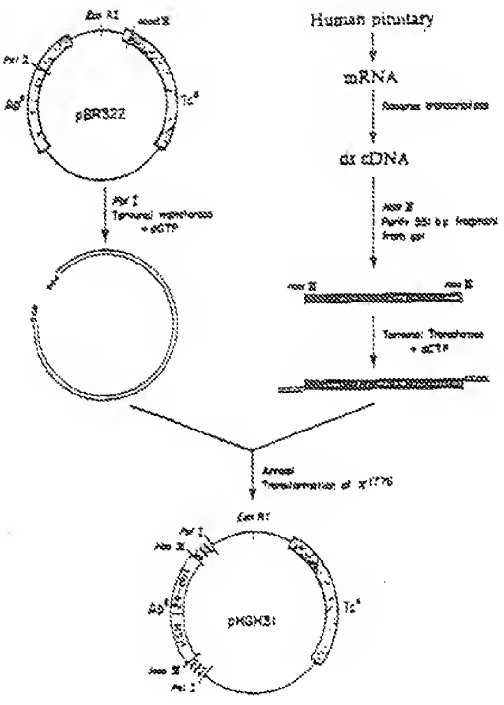


Fig. 3 Construction of a plasmid (pHGH31) containing coding sequences for amino acids 24–191 of human growth hormone. Poly(A)-mRNA was prepared from human pituitaries by a published procedure[17]. ds-cDNA (1.5 μg) was prepared from 5 μg of this RNA essentially as described by Wickens *et al.*[18], except that DNA polymerase 'Klenow fragment'[19] was substituted for DNA polymerase I in the second strand synthesis. cDNA (90 ng) was treated with *Hae*III, electrophoresed on an 8% polyacrylamide gel, and the region around 550 base pairs eluted. Approximately 1 ng of cDNA was obtained. Terminal deoxynucleotidyl transferase (TdT) was used to add approximately 20 dC residues per 3' terminus[19]. Annealing of the dC-tailed ds-cDNA with the dG-tailed vector DNA (60 ng) was performed in 150 μl of 10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.25 mM EDTA. The mixture was heated to 70 °C, allowed to cool slowly to 37 °C (12 h), then to 20 °C (6 h), before being used to transform χ1776 (ref. 20) by a published procedure[4].

ribosome binding site and initiation codon is 7 base pairs. The exact *lac* sequence in this region was restored as follows. The plasmid was cleaved with *Eco*RI, treated with S$_1$ nuclease to digest the single stranded ends, and reclosed by blunt-ended ligation. The plasmids from three individual transformants of χ1776 were analysed. In all instances the *Eco*RI site had been removed and a new *Alu*I site created in the plasmid (pHGH107–1).

5'...AGGAAACAG|CTATG...
3'...TCCTTTGTC|GATAC...

This structure, in which the ATG codon is now seven nucleotides downstream from the ribosome binding site, might be expected to increase the efficiency with which HGH mRNA is translated.

HGH radioimmune activity in bacterial extracts

Extracts of *E. coli* χ1776 containing the expression plasmids were tested for HGH expression by a direct radioimmunoassay (RIA) (Table 1). Easily detectable amounts of HGH were produced by bacteria containing pHGH107 or the derivative of pHGH107 lacking the *Eco*RI site (pHGH107–1). Contrary to

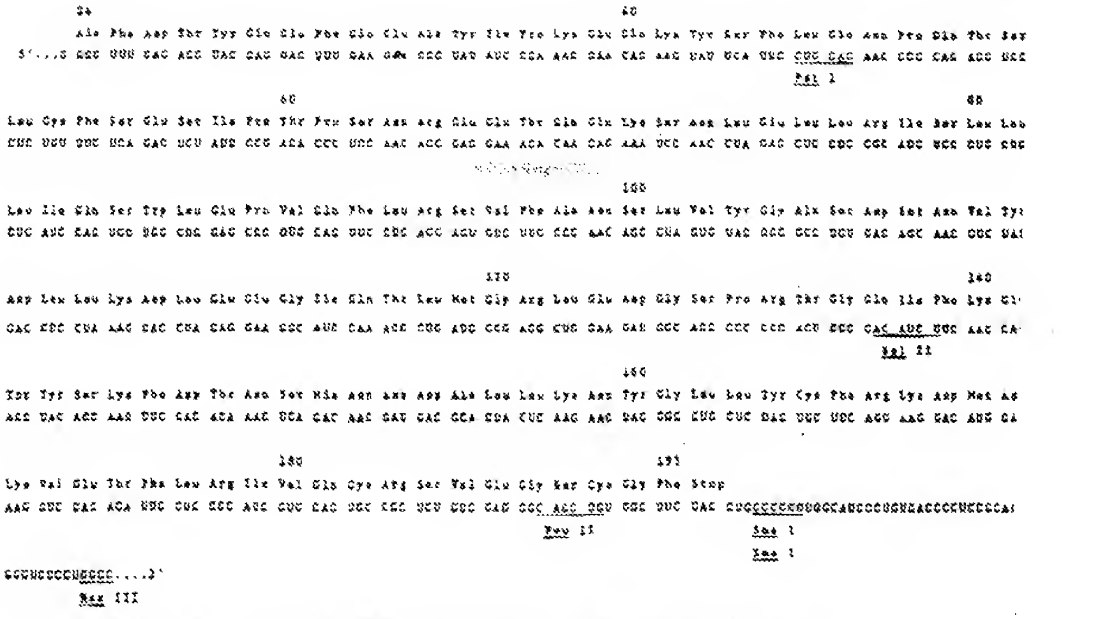


Fig. 3 The amino acid and mRNA sequences of HGH as determined by DNA sequencing[14] of pHGH31. The nucleotide sequence from the *Sma*I site to the *Hae*III site in the 3' noncoding region is from ref. 5, and was not determined here.

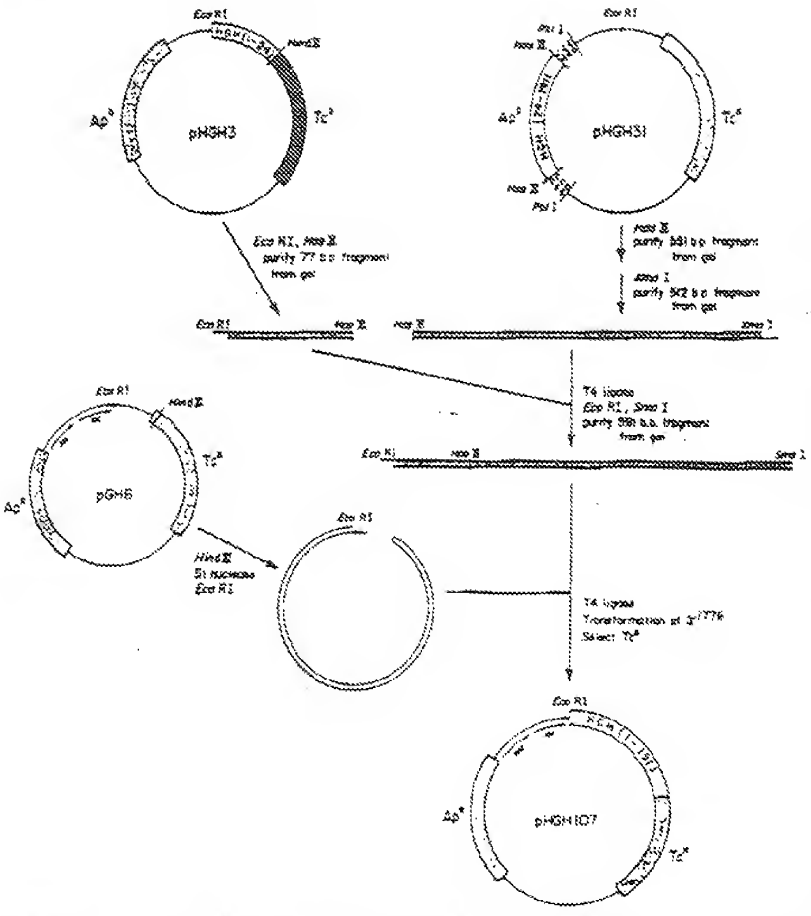


Fig. 4 Construction of a plasmid (pHGH107) for the bacterial expression of human growth hormone. pHGH3 (10 μg) was cleaved with *EcoRI* and *HaeIII* restriction endonucleases and the 77-base pair fragment containing coding sequences for HGH amino acids 1–23 was isolated from an 8% polyacrylamide gel. The plasmid pHGH31 (~5 μg) was cleaved with *HaeIII*. The 551-base pair HGH sequence and a co-migrating 540-base pair *HaeIII* fragment of pBR322 were purified by gel electrophoresis. Subsequent treatment with *XmaI* cleaved only the HGH sequence, removing 39 base pairs from the 5′ noncoding region. The resulting 512-base pair fragment was purified from the 540-base pair pBR322 *HaeIII* piece by electrophoresis on a 6% polyacrylamide gel. Samples of 0.3 μg of the 77-base pair *EcoRI–HaeIII* fragment and 0.3 μg of the 512-base pair *HaeIII–XmaI* fragment were polymerized with T_4 DNA ligase in a 16-μl reaction for 14 h at 4 °C. The mixture was heated at 70 °C for 5 min to inactivate the ligase, then treated with *EcoRI* (to cleave fragments which had dimerized through their *EcoRI* sites) and with *SmaI* (to cleave *XmaI* dimers), yielding a 591-base pair fragment with an *EcoRI* 'sticky' and and a *SmaI* 'blunt' end. After purification on a 6% polyacrylamide gel, approximately 30 ng of this fragment were obtained. The expression plasmid pGH6 containing tandem *lac* UV-5 promoters, was treated successively with *HindIII* nuclease S_1, and *EcoRI* and purified by gel electrophoresis. This removes most of the promoter sequence of the tc^R gene, but leaves the structural gene intact (25). 50 ng of the resulting vector, which had one *EcoRI* sticky end and one blunt end, was ligated to 10 ng of the 591-base pair HGH DNA. The ligation mixture was used to transform χ1776. Colonies were selected for growth on tetracycline (12.5 μg ml^{-1}). Since the tc^R promoter is no longer functional, tetracycline resistance is dependent on transcription from the *lac* promoters reading through the HGH gene sequence into the tc^R gene. Approximately 400 transformants were obtained. Colony screening by filter hybridisation[26] identified 12 colonies containing HGH sequences. The plasmids isolated from three of these colonies gave the restriction patterns expected for the correctly assembled gene when cleaved with *HaeIII*, *PvuII* and *PstI*. The DNA sequence of one clone, pHGH107, was determined.

our expectations, the introduction of four extra base pairs of DNA into the *lac* sequence in the region between the ribosome binding site and initiation codon results in increased HGH synthesis.

To verify that HGH expression is under *lac* operon control the plasmid pHGH107 was transformed into the *E. coli* K12 strain D1210, an I^Q derivative of *E. coli* K12 HB101 which overproduces *lac* repressor[32]. The data in Table 1 demonstrate that repressor effectively blocks HGH production and that this effect can be reversed by the addition of the inducer isopropyl-β-D-thiogalactoside.

Identification of HGH from bacterial extracts

Extracts of *E. coli* K12 strain RV308 (ref. 27) containing either pHGH107 or pBR322 were electrophoresed on SDS-polyacrylamide gels. The resulting protein patterns show one clear difference (Fig. 5a)—an additional protein, which co-migrates with HGH isolated from pituitaries, appears in the RV308/pHGH107 extract. To achieve partial purification of HGH, an extract of a 1-litre culture of χ1776/pHGH107 was prepared and fractionated as described in Fig. 5 legend. The polyacrylamide gel electrophoresis pattern of this partially

Table 1 Radioimmunoassay of HGH in extracts from bacteria containing expression plasmids

Strain	Cell density (cells per ml)	HGH by RIA (μg per ml)	HGH copies per cell
χ1776/pHGH107	3.69×10^8	2.4	186,000
	1×10^8	1.4	39,000
χ1776/pHGH107-1	3.6×10^8	1.3	116,000
	1×10^8	0.5	14,000
χ1776/pBR322	3.6×10^8	0	0
D1210/pHGH107	3.8×10^8	2×10^{-4}	15
D1210/pHGH107 (IPTG)	3.8×10^8	1.0	75,000

E. coli strains containing the appropriate plasmid were grown to the indicated cell density and collected by centrifugation. Stationary phase cultures of χ1776 contain 10^9 cells per ml. The cell pellet was resuspended in 55 μl of 50 mM Tris-HCl (pH 8), 50 mM EDTA, 15% sucrose, 1 mg ml^{-1} lysozyme, 0.025% lithium dodecylsulphate. After 30 min at 0 °C, 10 μl of 150 mM Tris-HCl (pH 7.5), 280 mM $MgCl_2$, 4 mM $CaCl_2$ and 1 μg of DNase I were added. The mixture was centrifuged for 15 min at 12,000*g*. Serial dilutions of the supernatants were analysed by radioimmunoassay using the Phadebas HGH PRIST kit (Pharmacia). Values are the average of three independent experiments expressed as μg per ml of cell culture. With HGH standard, the detection limit is 1 ml of culture was 0.02 ng. No activity was found in extracts of χ1776/pBR322. Where indicated, isopropyl-β-D-thiogalactoside (IPTG) was added to the cell culture at a concentration of 2 mM.

Fig. 5 Identification of HGH produced in bacteria by SDS-polyacrylamide gel electrophoresis. *a*, Protein patterns of crude extracts and partially purified HGH stained with Coomassie brilliant blue. Slot 1 contains 0.5 μg of pituitary HGH standard (Kabi), slot 2 contains a cell lysate of RV308/pHGH107, slot 3 contains a cell lysate of RV308/pBR322, slot 4 contains partially purified HGH isolated from χ1776/pHGH107. The samples were separated on a 15% polyacrylamide slab gel using the buffer system of Maizel[28] with the addition of 6 M urea. Crude lysates were prepared by growing cells in LB with 5 μg ml^{-1} tetracycline followed by lysis in 2% SDS, 1% β-mercaptoethanol. The lysates were precipitated with 10 volumes of cold acetone and the pellets were redissolved in SDS sample buffer for use in gel electrophoresis. The partially purified HGH was prepared from a stationary phase culture of χ1776/pHGH107. Cells were collected and resuspended in 1/50 of their original volume in 30 mM potassium phosphate (pH 7.0) containing 0.05 M NaCl and lysed by sonication. Polyethylenimine (Miles, Polymin-P) was added to 0.2%. After centrifugation for 1 h at 100,000*g*, ammonium sulphate was added to the supernatant to 60% saturation. The ammonium sulphate pellet was dissolved in 2 ml 10 mM potassium phosphate (pH 7.0), 0.3 M NaCl and chromatographed on a Sephacryl S-200 column (2.0 × 50 cm) equilibrated in the same buffer. The radioimmune active peak (assayed with a Pharmacia Phadebas HGH kit) was pooled, the protein concentrated by ammonium sulphate precipitation, the pellet redissolved in 1/10 volume of buffer, and the solution dialysed against 10 mM potassium phosphate (pH 7.0), 0.5 M NaCl. Precipitated material was removed by centrifugation resulting in a HGH preparation of approximately 25% purity. *b*, Autoradiograms of ^{35}S-labelled extracts of RV308/pHGH107. Slot 1 contains a total lysate of RV308/pHGH107 labelled with $H_2{}^{35}SO_4$. Slot 2 contains an ^{35}S-labelled extract of RV308/pHGH107 precipitated with α-HGH antiserum. The major (top) band co-migrated with unlabelled HGH standard (not shown). Cultures (1 ml) of RV308/pHGH107 were grown to A_{550} = 1 in low sulphur medium[29] containing 0.2 mCi ml^{-1} $H_2{}^{35}SO_4$, chased with 10 mM $MgSO_4$ for 5 min, collected and lysed using Triton X-100 and lysozyme[30]. Following DNase and RNase treatment[24], the lysate was mixed with a 10-fold excess of unlabelled RV308/pBR322 extract and diluted 1:1 into Triton immunoprecipitation buffer (0.15 M NaCl, 1% Triton X-100, 0.05 M Tris-HCl pH 7.5). Twenty μl of α-HGH antiserum (Kabi) was added per ml of original culture and the reaction was incubated for 12 h at 4 °C. The mixture was centrifuged and the supernatant incubated for 2 h with formaldehyde fixed *Staphylococcus* cells[31], filtered on 0.45-μm Nucleopore polycarbonate filters, washed with Triton immunoprecipitation buffer and extracted with SDS sample buffer[28]. The samples were run on a 15% slab gel containing urea and SDS as described above.

purified sample is shown in Fig. 5*a*, slot 4. The most prominent band co-migrates with HGH standard and its relative intensity is consistent with the 25% purity calculated from radioimmune activity.

In an independent attempt to confirm that the antigenically active component in the bacterial extracts co-migrates with HGH in SDS gels, we precipitated extracts from cells labelled with ^{35}S with antiserum prepared against HGH and separated the precipitated proteins by polyacrylamide gel electrophoresis. The autoradiogram of the immunoprecipitated material (Fig. 5*b*, slot 2) shows a major band, with electrophoretic mobility identical to that of HGH, as well as four smaller bands. All five of these bands are present in the unfractionated ^{35}S-labelled lysate (slot 1), whereas only the band with the same mobility as HGH can be seen in unlabelled extracts of cells grown in LB broth (Fig. 5*a*, slot 2). All five immune precipitated bands can be competed out by excess cold HGH (data not shown) and are thus antigenically related. This suggests that the four smaller HGH-related products observed in minimal media may be due to proteolytic degradation.

Additional evidence of proteolytic degradation of HGH is presented in Table 1. The levels of HGH in extracts from strains containing pHGH107 or pHGH107-1 were considerably higher in log-phase cultures than in stationary-phase cultures. Strains transformed with several other plasmids whose products are under *lac* promoter control show no comparable decrease of gene expression as a function of growth phase (unpublished results from this laboratory). It thus seems that there is some turnover of HGH in bacterial cells although we cannot rule out other explanations for the four smaller HGH-related products in the ^{35}S-labelled extracts such as premature termination of translation.

Despite the observed instability of HGH in certain growth conditions, the molecule seems to be relatively stable in log-phase cultures grown in rich media. In these conditions the amount of HGH in the cells (186,000 monomers per cell, Table 1) compares favourably with the expression level of other cloned genes using the same promoter in optimised conditions[32].

Conclusion

Using a novel combination of chemically synthesised DNA and cDNA, a recombinant *E. coli* strain has been constructed which produces HGH in large amounts. This is the first time that a human polypeptide has been directly expressed in *E. coli* in a non-precursor form. The hybrid DNA cloning techniques described as a route to the cloning and expression of HGH coding sequences in *E. coli* are generally applicable to other polypeptides which are synthesised initially as inactive precursors and later processed, or for which full length cDNA transcripts are unavailable.

We thank John Shine for help in planning the project, Elizabeth McLeod for assistance, Ron Wetzel for help in peptide identification, Frances Fields and Dale Cullins for technical support, Alane Gray for preparing illustrations, and Robert Swanson and Herbert Boyer for their encouragement. RNA used was prepared by P.H.S. while a postdoctoral fellow in the Department of Biochemistry and Biophysics, University of California at San Francisco. Additional RNA was obtained from pituitaries donated by Drs Gerald Silverberg and O. H. Pearson. This work was supported by a contract from KabiGen AB. All experiments involving recombinant genomes were performed in accordance with current NIH guidelines.

Received 4 July; accepted 16 August 1979.

1. Raben, M. S. *J. clin. Endocr.* 18, 901–904 (1958).
2. *Proc. of NIAMDD Symp.* Publ. No. 74-612 (ed. Raiti, S.) (Baltimore, Maryland, 1973).
3. Bancroft, F. C. *Expl Cell Res.* 79, 275–278 (1973).
4. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. *Nature* 270, 486–494 (1977).
5. Goodman, H. M. *et al.* in *Specific Eukaryotic Genes* (eds Engberg, J., Klenow, H. & Leick, V.) 179–190 (Munksgaard, Copenhagen, 1979).
6. Villa-Komaroff, L. *et al. Proc. natn. Acad. Sci. U.S.A.* 75, 3727–3731 (1978).
7. Seeburg, P. H. *et al. Nature* 276, 795–798 (1978).
8. Mercereau-Puijalon, O. *et al. Nature* 275, 505–510 (1978).
9. Fraser, T. H. & Bruce, B. J. *Proc. natn. Acad. Sci. U.S.A.* 75, 5936–5940 (1978).
10. Chang, A. C. Y. *et al. Nature* 275, 617–624 (1978).
11. Itakura, K. *et al. Science* 198, 1056–1063 (1977).
12. Goeddel, D. V. *et al. Proc. natn. Acad. Sci. U.S.A.* 76, 106–110 (1979).
13. Crea, R., Kraszewski, A., Hirose, T. & Itakura, K. *Proc. natn. Acad. Sci. U.S.A.* 75, 5765–5769 (1978).
14. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* 74, 560–564 (1977).
15. Bolivar, F. *et al. Gene* 2, 95–113 (1977).
16. Backman, K., Ptashne, M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* 73, 4174–4178 (1976).
17. Ullrich, A. *et al. Science* 196, 1313–1319 (1977).
18. Wickens, M. P., Buell, G. N. & Schimke, R. T. *J. biol. Chem.* 253, 2483–2495 (1978).
19. Klenow, H. & Henningsen, I. *Proc. natn. Acad. Sci. U.S.A.* 65, 168–175 (1970).
20. Curtiss, R. III *et al.* in *Molecular Cloning of Recombinant DNA* (eds Scott, W. A. & Werner, R.) 99–111 (Academic, New York, 1977).
21. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* 72, 3961–3965 (1975).
22. Shine, J., Seeburg, P. H., Martial, J. A., Baxter, J. D. & Goodman, H. M. *Nature* 270, 494–499 (1977).
23. Li, C. H. *Proc. Am. phil. Soc.* 116, 365 (1972).
24. Backman, K. & Ptashne, M. *Cell* 13, 65–71 (1978).
25. Boyer, H. W. *et al.* in *Recombinant Molecules: Impact on Science and Society* (eds Beers, R. F. & Bassett, E. G.) 9–20 (Raven, New York, 1977).
26. Maizels, N. *Nature* 249, 647 (1974).
27. Maurer, R. thesis, Harvard Univ. (1978).
28. Maizel, J. V. Jr *Meth. Virol.* 5, 180–246 (1971).
29. Roberts, J. & Roberts, C. *Proc. natn. Acad. Sci. U.S.A.* 72, 147–151 (1975).
30. Katz, L., Kingsbury, D. T. & Helinski, D. R. *J. Bact.* 114, 577–591 (1973).
31. Kessler, S. W. *J. Immun.* 115, 1482–1490 (1975).
32. Roberts, T. M., Kacich, R. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* 76, 760–764 (1979).

90. *Construction and Identification of a Bacterial Plasmid Containing the Human Fibroblast Interferon Gene Sequence*

By Tadatsugu TANIGUCHI,*) Masaharu SAKAI,*)
Yoshiaki FUJII-KURIYAMA,*) Masami MURAMATSU,*)
Shigeyasu KOBAYASHI,**) and Tetsuo SUDO***)

(Communicated by Toshio KUROKAWA, M.J.A., Nov. 12, 1979)

Introduction. Interferons are species specific glycoproteins (Weil and Dorner, 1973) which are produced by various cells upon induction with viruses and double-stranded RNAs as well as other compounds, imparting an antiviral state in the recipient cells (Isaacs and Lindenmann, 1957). Beside their antiviral function, other multiple functions such as antitumor and anticellular functions have been described (see The Interferon System, 1977 for reviews). Interferons produced in cultures of human buffy coat cells and in human fibroblasts exhibit a number of distinct properties (Havell *et al.*, 1975; Berg *et al.*, 1975; Gresser *et al.*, 1974; Vilcek *et al.*, 1977), suggesting that leucocyte interferon and fibroblast interferon are different proteins coded for by distinct structural genes (Cavalieri *et al.*, 1977).

A number of attempts have been made for physical and chemical characterization of the interferon molecule and for its clinical investigation. However, mainly due to the low amount of this protein produced by the cell, it is as yet difficult to obtain sufficient quantities of purified interferon for these studies.

One approach to solve the problem appears to be the use of recombinant DNA technology. As has been the case for many proteins, cloning of the gene coding for interferon will provide a definitive answer to the molecular structure of this protein as well as the organization of the gene. In addition, with the aid of gene manipulation, cloned DNA may serve as a potent interferon producer in bacterial or in cognate host. Here, we present evidence for successful construction of a hybrid plasmid DNA containing a human fibroblast interferon gene sequence.

Materials and methods. Human foreskin fibroblast strain DIP2, a high interferon producer, has been isolated by us. Superinduction

*) Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Tokyo.
**) Basic Research Laboratories, Toray Industries, Inc., Kamakura.
***) The Tokyo Metropolitan Institute of Medical Science, Tokyo.

of the cells by Poly (I) : (C) and cycloheximide, and isolation of total cytoplasmic mRNA were carried out essentially as described (Sehgal *et al.*, 1978). Synthesis of double-stranded cDNA (Maniatis *et al.*, 1976), transformation of *Escherichia coli* χ 1776 by pBR322-cDNA hybrid (Enea *et al.*, 1975) and detection of the specific recombinant DNA by *in situ* colony hybridization (Grunstein and Hogness, 1975) were performed essentially according to the published procedures. Preparation of plasmid DNA was according to the procedure of Currier and Nester (1976) with some modifications. Trapping of the mRNA-DNA hybrid on nitrocellulose filter (Nygaard and Hall, 1963) and elution of the mRNA from the filter (Harpold *et al.*, 1978) were carried out essentially as described. Microinjection of mRNA into *Xenopus* oocytes was performed according to Gurdon *et al.* (1971). The interferon activity was determined following the procedure of Suzuki *et al.* (1974).

Results and discussion. Total cytoplasmic RNA was extracted from human fibroblast DIP2 cells after four hours induction by poly (I) : (C) in the presence of cycloheximide and poly(A) containing mRNA was isolated by oligo(dT)-cellulose affinity chromatography. Starting from 1.5×10^9 cells, 250 μg of poly(A) containing mRNA was obtained. After fractionation of the mRNA with a 5%-25% sucrose gradient centrifugation into 20 fractions, small portion of the mRNA from the fractions around 12S region was injected into *Xenopus* oocytes and interferon activity was determined as described in Materials and methods.

Table I. Interferon mRNA activity in the fractions from sucrose gradient

Fraction No.	Interferon activity (units/ml)
9	<50
10	44
11	550
12	52

RNA from sucrose gradient (input 250 μg) was precipitated, dissolved in 20 μl H_2O and an aliquot (about 0.2 μl) was injected into *Xenopus* oocytes as described (Gurdon *et al.*, 1971). Interferon activity was measured according to Suzuki *et al.* (1974).

As shown in Table I, the highest interferon mRNA activity was detected in fraction 11. This fraction contained about 5 μg of mRNA (termed "interferon mRNA" here), which was used as the template for cDNA synthesis. About 1.5 μg of double-stranded cDNA was synthesized, elongated with dAMP residues, then hybridized with EcoRI-cleaved, dTMP-elongated pBR322. This hybrid DNA was used

to transform *E. coli* strain χ 1776 in the P3 laboratory of Cancer Institute and colonies containing the hybrid plasmid DNA were selected on agar plates containing ampicillin (20 μg/ml). Efficiency of the transformation was about 1.5×10^5 colony/μg cDNA. For the first screening, 3600 colonies were picked up and transferred on grid-meshed nitrocellulose filters in triplicate. After the colonies had grown on the filters, DNA from duplicate filters was fixed for *in situ* colony hybridization (Grunstein and Hogness, 1975). Colonies grown on the third filter were kept at 4°C.

For colony hybridization, two kinds of ^{32}P-labeled cDNA probe were prepared as follows. First, partially purified "interferon mRNA" (6 μg) was prepared as described above and ^{32}P-labeled cDNA synthesized (0.45 μg, specific radio activity: 6×10^8 c.p.m./μg). After removing the template RNA by alkali treatment, cDNA was hybridized with about 50 fold excess of mRNA prepared from mock-induced cells incubated for four hours in the presence of cycloheximide. Non-hybridized cDNA consisting of about 10% of the total radioactivity was separated from mRNA-cDNA hybrid by hydroxyapatite column chromatography and this cDNA was used as probe A. After alkali treatment to remove mRNA, cDNA which had hybridized with mRNA was used as probe B. Both probe A and probe B were separately hybridized with colony DNA fixed on nitrocellulose filters (Grunstein and Hogness, 1975). Colonies hybridized with probe A but failed to hybridize or hybridized much less with probe B were screened by autoradiogram. Four colonies, nos. 319, 644, 746 and 3578, were in this category (Table II).

Table II. Colony hybridization of the cDNA clones with different cDNA probes

Ampicillin resistant colony	Extent of hybridization	
	probe A	probe B
no. 319	+++++	++
no. 644	+++	+
no. 746	++	—
no. 3578	+++++	+

Colony hybridization was done according to Grunstein and Hogness (1975). Extent of hybridization was determined by visual inspection of the autoradiogram.

For the second screening, these four colonies were grown in liquid media and each recombinant plasmid DNA was prepared as described in Materials and methods. 5 μg of each recombinant plasmid or pBR322 DNA was linealized by Hind III digestion, denatured, and then hybridized with the "interferon mRNA" (2.5 μg) under the con-

ditions where RNA-DNA hybrid could be formed whereas DNA renaturation was negligible (Casey and Davidson, 1977; details to be published elsewhere). After 4 hrs of hybridization, single-stranded plasmid DNA whose cDNA part should have been hybridized with complementary mRNA was trapped on nitrocellulose filter (Nygaard and Hall, 1963) and hybridized mRNA was eluted from the filter (Harpold *et al.*, 1978). After oligo (dT)-cellulose column chromatography to remove possible contaminants such as DNA, mRNA was microinjected into frog oocytes to determine the interferon mRNA activity. As shown in Table III, mRNA hybridized with the recombinant plasmid DNA from clone no. 319 gave rise to interferon synthesis whereas mRNA hybridized with the other plasmid DNA including pBR322 failed to synthesize interferon in frog oocytes. When the same amount of "interferon mRNA" as that used for hybridization was incubated as a control under the same conditions without DNA, chromatographed on the oligo(dT)-cellulose column and injected into oocytes, approximately 1000 interferon units per ml equivalent was synthesized. Since about 50% of the DNA (i.e. mRNA-DNA hybrid) can be trapped on the filter under our conditions, interferon mRNA must have hybridized with clone no. 319 DNA at an efficiency of about 72% of the input or even higher (Table III).

Table III. Hybridization-translation assay with interferon mRNA and various recombinant DNA

Plasmid DNA	Interferon activity (units/ml)
from no. 319	360
from no. 644	<10
from no. 746	15
from no. 3578	<10
pBR322	<10

"Interferon mRNA" was hybridized with linealized, denatured plasmid DNA and the hybrid was trapped on nitrocellulose filter. RNA was eluted, chromatographed on an oligo (dT)-cellulose column, then injected into frog oocytes for interferon synthesis. Interferon activity was measured as described in Meterials and methods.

In the next experiment, double amount of the mRNA and plasmid DNA from clone no. 319 as those used in the previous experiment was mixed in the hybridization solution and a half portion of the mixture was immediately withdrawn for filter trapping (zero time) and the rest was allowed to hybridize for 4 hrs. Interferon mRNA activity in the RNA trapped on the filter was determined as described above. The interferon activity at zero time and after 4 hrs incubation was

<30 units/ml and 150 units/ml, respectively.

These results thus indicate that the recombinant plasmid DNA from clone no. 319, named TpIF319, contains the sequence for the interferon mRNA. We have also observed that when TpIF319 DNA was hybridized with a mixture of "interferon mRNA" and rabbit globin mRNA, interferon mRNA activity but not globin mRNA activity was detectable from the filter-trapped RNA, whereas an opposite result was obtained when TpIF319 DNA was replaced with pBR322 DNA containing rabbit β-globin gene insert. In addition, both mRNAs were trapped on the filter when two of the recombinant plasmid DNA were present in the hybridization mixture (to be published elsewhere).

In conclusion, we have repotred in this paper the construction of the recombinant plasmid DNA which contains human fibroblast interferon gene sequence. The method for the identification of the specific recombinant which we employed in this paper should be applicable for the cloning of any kind of gene which is expressed in even smaller amount in the cell, provided that the mRNA translation product can be identified. The recombinant plasmid we have constructed will be useful not only for defining the hitherto unknown structure of the interferon gene but also for the attempts of the mass production of this protein.

Acknowledgement. We wish to express our sincere gratitude to Dr. H. Sugano, Director of Cancer Institute, for his continuous support and encouragement. Thanks are also due to Drs. Y. Mishima, R. Kominami and H. Koyama, Cancer Institute, for valuable suggestions and discussions. Skilfull technical assistance of Ms. M. Ohkawa is also acknowledged.

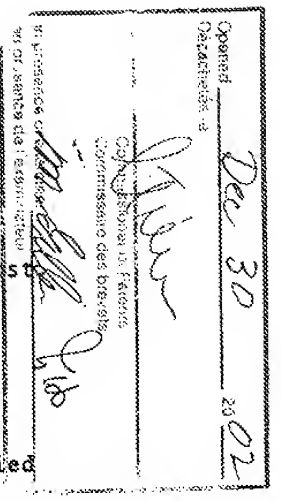
References

Baron, S., and Dianzani, F. (eds.) (1977): The Interferon System. Tex. Rep. Biol. Med., 35, 1–573.
Berg, K., Ogburn, C. A., Paucker, K., Mogensen, K. E., and Cantell, K. (1975): J. Immunol., 114, 640–644.
Casey, J., and Davidson, N. (1977): Nucleic Acids Res., 4, 1539–1552.
Cavalieri, R. L., Havell, E. A., Vilček, J., and Pestka, S. (1977): Proc. Natl. Acad. Sci. USA, 72, 3287–3291.
Currier, T. C., and Nester, E. W. (1976): Analyt. Biochem., 76, 431–441.
Enea, V., Vovis, G. F., and Zinder, N. D. (1975): J. Mol. Biol., 96, 495–509.
Gresser, I., Bandu, M. T., Brouty-Boye, D., and Tovey, M. (1974): Nature, 251, 543–545.
Grunstein, M., and Hogness, D. S. (1975): Proc. Natl. Acad. Sci. USA, 72, 3961–3965.
Gurdon, J. B., Lane, C. D., Woodlans, H., and Marbaix, G. (1971): Nature, 233, 177–179.

Harpold, M. M., Dobner, P. R., Evans, R. M., and Bancroft, F. C. (1978): Nucleic Acids Res., 5, 2039–2053.
Havell, E. A., Berman, B., Ogburn, C. A., Berg, K., Paucker, K., and Vilček, J. (1975): Proc. Natl. Acad. Sci. USA, 72, 2185–2187.
Isaacs, A., and Lindenmann, J. (1957): Proc. R. Soc. London, ser. B., 147, 258–267.
Maniatis, T., Sim, G. K., Efstratiadis, A., and Kafatos, F. C. (1976): Cell, 8, 163–182.
Nygaard, A. P., and Hall, B. D. (1963): Biochem. Biophys. Res. Commun., 12, 98–104.
Sehgal, P. B., Lyles, D. S., and Tamm, I. (1978): Virology, 89, 186–198.
Suzuki, J., Akaboshi, T., and Kobayashi, S. (1974): Japan. J. Microbiol., 18, 449–456.
Vilček, J., Havell, E. A., and Yamazaki, S. (1977): Ann. N. Y. Acad. Sin., 284, 705–710.
Weil, R., and Dorner, F. (1973): Selective Inhibitors of Viral Functions (ed. Carter, W. A.), CRC Press, Cleaveland, Ohio, pp. 107–121.

Mailing List for "Interferon" Manuscripts

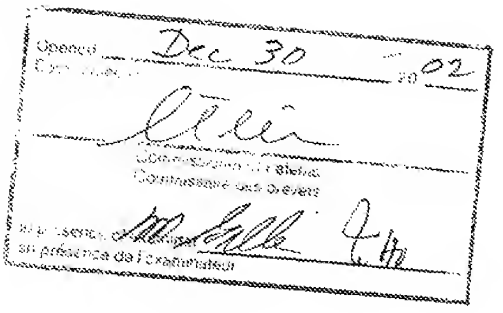
Aebi, Bern
Aloni, Rehov
Anderson, Bethesda
Arber, Basel
Axel, Columbia
August, Bronx
Baltimore, Cambridge
Baron, Galveston
Bautz, Heidelberg
Bayev, Moscow
Berg, Palo Alto
Bernardi, Paris
Birnstiel, Zürich
Bishop, San Francisco
Biogen, Geneva
Borst, Amsterdam
Bové, Versailles
Brenner, Cambridge
Bres, Geneva
Bretscher, Cambridge
Brown, Baltimore
Burton, Oxford
Cantoni, Bethesda
Cantell, Helsinki
Carbon, Santa Barbara
Carter, Buffalo
Cawthorn, Geneva
Chamberlin, Berkeley
Chambon, Strasbourg
Chany, Paris (& Villejuif)
Clark, Cambridge
Clercq, Leuven
Coffin, Boston
Cohen, Stanford
Cohn, San Diego
Crick, San Diego
Cross, Wellcome
Curtis, Philadelphia
Curtiss III, Birmingham
Dahlberg, A.E., Bethesda
Dahlberg, J.E., Madison
Darnell, Bronx
Davies, Geneva
Davis, Stanford
de Maeyer, Orsay
Desmyter, Leuven
Doty, Cambridge
Duesberg, Berkeley
van der Eb, Leiden
Echols, Berkeley
Efstratiadis, Cambridge
Epstein, Geneva
Falcoff, Paris
Feix, Freiburg
Fiers, Gent
Finter, Beckenham
Flavell, London
Forget, Boston
Fresco, Princeton
Fraenkel-Conrat, Berkeley
Gallo, Bethesda
Garen, New Haven
Gesteland, Cold Spring Harbor
Gierer, Tübingen
Gilbert, Cambridge
Gilham, Lafayette
Gilvarg, Princeton
Glaser, Berkeley
Goodman, San Francisco
Gresser, Villejuif
Gros, Paris
Gruber, Groningen
Grunberg-Manago, Paris
Hailey, Fish & Neane
Hanafusa, New York
Harris, Adelaide, S. Aust
Hartley, London
Haselkorn, Chicago
Havell, New York
Hindley, Bristol
Hinnen, Basel
Hogness, Boston
Hofman, Zürich
Hofschneider, Martinsried
Hohn, Basel
Holley, La Jolla
Hood, Pasadena
Horecker, Nutley
Hurwitz, Bronx
Jacob, Paris
Jeffreys, Leicester
Kaesberg, Madison
Kafatos, Cold Spring Harbor
Kamen, London
Kaziro, Tokyo
Kellenberger, Basel
Kerr, London
Khorana, Cambridge
Klingmüller, München
Knight, Wilmington
Konigsberg, New Haven
Koprowski, Philadelphia
Kornberg, Palo Alto
Kourilsky, Paris
Krakow, New York
Lawrason, Kenilworth
Leder, Bethesda
Lederberg, Stanford
Leibowitz, Bloomfield
Lengyel, New Haven
Levy, Bethesda
Libonati, Napoli
Lindenmann, Zürich
Lingrel, Cincinnati
Littauer, Rehovoth
Lodish, Cambridge
Luria, Cambridge
Lüscher, Bern
Mach, Geneva
Maniatis, Pasadena
Marks, New York
Marmur, Bronx
Martin, Bethesda
Merigan, Stanford
Murray, Heidelberg
Montagnier, Paris
Nathans, Baltimore
Nirenberg, Bethesda



This is EXHIBIT FIERS-39
to
the Affidavit of Walter C. Fiers
sworn before me
this 29 th day of November, 2001

Commissioner for Oath or Notary Public

Ochoa, Nut
Oeschger, Baltimore
O'Malley, Houston
Parsons, Charlottesville
Paul, Glasgow
Pitha, Baltimore
Pestka, Nutley
Ptashne, Cambridge
Rajbhandari, Madison
Rauscher, Bethesda
Reich, New York
Revel, Rehovoth
Rich, Cambridge
Ross, Madison
Rowe, Bethesda
Ruddle, New Haven
Sanger, Cambridge
Schaller, Heidelberg
Schwartz, New York
Shapiro, Bronx
Sharp, Cambridge
Singer, Bethesda
Skalka, Nutley
Sinsheimer, Pasadena
Smith, Cambridge
Spahr, Geneva
Speyer, Storrs
Spiegelman, New York
Sgaramella, Pavia
Staehelin, Basel
Steitz, New Haven
Stent, Berkeley
Stewart, New York
Strander, Stockholm
Streisinger, Eugene
Stutz, Neuchâtel
Subak-Sharpe, Glasgow
Swetley, Vienna
Szybalski, Madison
Taniguchi, Tokyo
Temin, Madison
Tener, Vancouver
Thomas, Boston
Tissières, Geneva
Tocchini-Valentini, Napoli
Tonegawa, Basel
Tooze, Heidelberg
Veno, Düsseldorf
Vilček, New York
Viñuela, Madrid
Vogt, Los Angeles
Walker, Edinburgh
Warner, Irvine
Watson, Cold Spring Harbor
Weber, R., Bern
Weissman, New Haven
Williamson, London
Wittmann, Berlin
Yanofsky, Stanford
Zachau, München
Zillig, München
Zinder, New York



Human leukocyte and fibroblast interferons are structurally related.

Tadatsugu Taniguchi*, Ned Mantei°, Marco Schwarzstein°, Shigekazu Nagata°, Masami Muramatsu* and Charles Weissmann°

°Institut für Molekularbiologie I, Universität Zürich, 8093 Zürich, Switzerland

*Dept. of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Tokyo 170, Japan

This is EXHIBIT FIERS-40
to
the Affidavit of Walter C. Fiers
sworn before me
this 13 th day of November, 2001

Commissioner for Oath or Notary Public

SUMMARY

The coding sequences of the cDNAs of cloned human leukocyte interferon I and human fibroblast interferon show homologies of 45% at the nucleotide and 29% at the amino acid level. We estimate that the two genes were derived from a common ancestor about 500 to 1'000 million years ago.

The acid-stable human interferons are subdivided into two major groups, namely fibroblast interferons (F-IF) and leukocyte interferons (Le-IF); these are the major components of the interferons produced by induced fibroblasts and leukocytes, respectively. Some cells, such as the lymphoblastoid Namalva cell line, produce a mixture of 90% Le-IF and 10% F-IF (1, 2). The two interferon types have several features in common: both are glycoproteins with molecular weights ranging from 16'000 to 26'000 (3-9), the induction and shut-off of their synthesis appears to be under similar control (6), and at least some of the responses elicited in target cells are similar, such as induction of an antiviral state, which is accompanied by increased synthesis of several proteins (10-13). Nonetheless, the two kinds of interferons differ in many respects. Antibodies directed against Le-IF do not neutralize F-IF and vice-versa (14), the target cell specificities of the two IFs differ (15), and the sequences of the 13 amino terminal amino acids of F-IF and of Le-IF (from lymphoblastoid cells) show no homology (16, 22). Although Le-IF and F-IF are encoded by different mRNA species (17), it is not known whether these mRNAs are transcribed from distinct genes or whether they arise from the same gene via a common precursor which is processed or spliced in different modes.

We have recently cloned and sequenced one species each of Le-IF (Le-IF I) (18,19) and F-IF cDNA (20,21). A second

species of Le-IF (Le-IF II) cDNA has recently been identified (M. Streuli, S. Nagata and C. Weissmann, unpublished results).

In Fig. 1 the nucleotide sequences of Le-IF I and F-IF cDNA were aligned so that the AUGs closest to their 5' termini coincided. From the amino terminal sequence published for F-IF (16) and lymphoblastoid Le-IF (22) one can determine that in the case of F-IF the 21st codon following the initiation triplet and in the case of Le-IF the 23rd codon represents the first amino acid of the interferon polypeptide. Presumably the stretch in between encodes a signal peptide. Since the putative signal peptide of Le-IF comprises 23 and that of F-IF 21 amino acids, the IF polypeptides, as aligned in Fig. 1, are shifted by two residues relative to their termini. In this alignment, 48 of 166 positions (29%) of the interferon polypeptides have identical amino acids. To plot the degree of homology between the F-IF and Le-IF as function of the map distance, the sequence was subdivided into segments of 8 amino acids (or 24 nucleotides), each overlapping by 4 amino acids (or 12 nucleotides) with the neighboring segments, and the percent coincidence of amino acids (and nucleotides) for each segment was determined (cf. van Ooyen et al., ref. 23). As seen in Fig. 2, amino acid sequences show three domains of homology. The first one, with the least degree of homology, corresponds to the putative signal sequence, which is rich in hydrophobic residues and has 4 identical amino acid positions out of 21; the second domain, between the 28th and 80th amino acid (counted on the Le-IF sequence), has 21 identical residues out of 51 (41% homology) and the third, between

positions 115 and 151 (Le-IF sequence), has 19 out of 35 identical residues (54%). The longest stretches of contiguous conserved amino acids are Gln-Phe-Gln-Lys (pos. 47-50 of Le-IF and 49-52 of F-IF) and Cys-Ala-Trp (pos. 139-141 and pos. 141-143, respective the latter sequence is notable because it comprises Cys and Trp, which are preferentially conserved in related proteins (24). Table 1 shows that conservation was highest between the interferon polypeptides (not considering the signal sequences) for Trp, Phe, Arg, Cys and Tyr residues, in agreement with the general experience that the amino acids most likely to be conserved between related proteins are Trp > Cys > Tyr > Arg > Phe, His (24). Even where amino acids are conserved, the codons show one or more nucleotide changes in half the instances. The codons of 3 out of 7 conserved Leu residues are non-related, as are 2 of 4 codons pertaining to conserved Ser residues. This suggests that there is a strong selective pressure favoring the conservation of several amino acids. It is quite likely that at least some of the conserved amino acids are essential for a function common to Le-IF and F-IF, perhaps the induction of the virus-resistant state in the target cell. These findings may provide guidelines for the tailoring of modified (25), possibly shorter polypeptides possessing certain activities of interferon.

The nucleic acid sequences show an average homology of 43% in the domain of the signal sequence and of 45% in the interferon polypeptide sequence. On a random basis, about 25% of the nucleotide positions should

coincide. Within the interferon coding sequence, the nucleotide homologies are more evenly distributed than the amino acid homologies. However, one may distinguish, albeit to a less pronounced degree, the same two blocks of similarity noted for the amino acids. The longest region without mismatches extends for 13 nucleotides (cf. 47th to 51st codon of Le-IF vs. 49th to 53rd codon of F-IF). There are, in addition, sequences of 17, 18 and 20 nucleotides with 3, 3 and 4 mismatches, respectively. The heteropolymeric 3'-terminal non-coding region of Le-IF cDNA has 242 nucleotides, and is longer by 39 residues than its counterpart in F-IF cDNA. In aligning the two sequences four gaps were introduced to maximize homology, as described by van Ooyen et al.[23] Thereby, several segments were matched with 29 to 41% homology. The introduction of gaps in the alignment may be justified in view of the arguments presented previously, that introns and non-coding regions of reduplicated genes diverge as a consequence of block insertions and/or deletions in the course of evolution (23,26).

It is unlikely that the extent of homology between Le and F-IF cDNA would allow meaningful crosshybridization between the two species.

On the basis of our findings there is no doubt that Le-IF and F-IF genes are derived from a common ancestral sequence. When did the separation of these genes occur? Human α and β globin show 57% amino acid mismatches, and human β-globin and myoglobin, as well as α-globin and myoglobin, 76% mismatches.

If the rate of divergence of interferons and globins is comparable (however, cf. p. 50, ref. 24, for proteins showing both higher and lower rates) then the separation of interferon genes occurred after that of myoglobin and hemoglobins and before that of α- and β-globins, i.e. between 500 and 1000 million years ago (24). The interferon genes may thus be about as old as the vertebrates (27).

ACKNOWLEDGEMENTS

T.T. and M.M. are indebted to Dr. H. Sugano for his continuous support and interest. The work at the University of Zürich was supported by Biogen S.A. and the Schweizerische Nationalfonds.

REFERENCES

1. Havell, E.A., Yip, Y.K. & Vilček, J. J. gen. Virol. 38, 51-59 (1977).

2. Paucker, K. The Interferon System, Texas Reports on Biology and Medicine (Baron, S. & Dianzani, F., eds.) 35, 23-28 (1977).

3. Bridgen, P.J., Anfinsen, C.B., Corley, L., Bose, S., Zoon, K.C., RÛegg, U.T. & Buckler, C.E. J. Biol. Chem. 252, 6585-6587 (1977).

4. Knight, E., jr. Proc. Natl. Acad. Sci. USA 73, 520-523 (1976).

5. Tan et al., J. Biol. Chem. 254, 8067-8073 (1979).

6. Hayes, T.G., Yip, Y.K. & Vilček, J. Virology 98, 351-363 (1979).

7. Rubinstein, M., Rubinstein, S., Familletti, P.C., Miller, R.S., Waldman, A.A. & Pestka, S. Proc. Natl. Acad. Sci. USA 76, 640-644 (1979).

8. Stewart, W.E., II, & Wiranowska-Stewart, M. Virology 97, 473-476 (1979).

9. Havell, E.A., Yamazaki, S. & Vilček, J. J. Biol. Chem. 252, 4425-4427 (1977).

10. Lebleu, B., Sen, G.C., Shaila, S., Cabrer, B. & Lengyel, P. Proc. Natl. Acad. Sci. USA 73, 3107-3111 (1976).

11. Hovanessian, A.G., Brown, R.E. & Kerr, I.M. Nature 268, 537-540 (1977).

12. Hovanessian, A.G. & Kerr, I.M. Eur. J. Biochem. 93, 515-526 (1979).

13. Schmidt, A., Chernajovsky, Y., Shulman, L., Federman, P., Berissi, H. & Revel, M. Proc. Natl. Acad. Sci. USA 76, 4788-4792 (1979).

14. Havell, E.A., Berman, B., Ogburn, C.A., Berg, K., Paucker, K. & Vilček, J. Proc. Natl. Acad. Sci. USA 72, 2185-2187 (1975).

15. Stewart, W.E., II, In "The Interferon System" (Springer-Verlag, Wien & New York), pp. 134-145 (1979).

16. Knight, E., jr., Hunkapiller, M.W., Korant, B.D., Hardy, R.W.F. & Hood, L.E. Science 207, 525-526 (1980).

17. Cavalieri, R.L., Havell, E.A., Vilček, J. & Pestka, S. Proc. Natl. Acad. Sci. USA 74, 3287-3291 (1977).

18. Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Ecsödi, J., Boll, W., Cantell, K. & Weissmann, C. Nature, in press (1980).

19. Mantei, N., Schwarzstein, M., Streuli, M., Panem, S., Nagata, S. & Weissmann, C., Gene, in press (1980).

20. Taniguchi, T., Sakai, M., Fujii-Kuriyama,Y., Muramatsu, M., Kobayashi, S. & Sudo, T. Proc. Jap. Acad. Ser. B, 55, 464-469 (1979).

21. Taniguchi, T., Ohno, S., Fujii-Kuriyama, Y. & Muramatsu, M. Gene, in press (1980).

22. Zoon, K.C., Smith, M.E., Bridgen, P.J., Anfinsen, C.B., Hunkapiller, M.W. & Hood, L.E. Science 207, 527-528 (1980).

23. van Ooyen, A., van den Berg, J., Mantei, N. & Weissmann, C. Science 206, 337-344 (1979).

24. Dayhoff, M.O., In "Atlas of Protein Sequence and Structure" (Natl. Biomedical Res. Found., Washington, D.C.) Vol. V (1972)

25. Müller, W., Weber, H., Meyer, F. & Weissmann, C. J. Mol. Biol 124, 343-358 (1978).

26. Konkel, D.A., Maizel, J.V., jr. & Leder, P. Cell 18, 865-873 (1979).

27. Kurtén, B. In "Ideas in modern biology" Proceedings vol. 6, XVI International Congress of Zoology (Moore, J.A., ed.), Natural History Press, N.Y. (1965) pp. 327-354.

TABLE 1 Conservation of amino acids in leukocyte and fibroblast interferon.*

	F-IF	Le-IF	Conserved amino acids	Number of changes in codon of conserved amino acids 0	1	2	3
Leu	25	22	8	1	4	3	
Cys	3	5	2	1	1		
AsN	12	6	1	1			
Arg	11	12	5	1	3	1	
Phe	9	8	4	2	2		
Pro	1	6	1		1		
GlN	11	10	3	3			
Lys	11	8	3	2	1		
Ala	6	10	2	2			
Glu	13	15	4	4			
Ile	11	7	3	2	1		
Ser	9	13	4		2	1	1
Trp	3	2	2	2			
Tyr	10	4	4	1	3		
Val	5	6	1	1			
Asp	5	11	1	1			
Thr	6	9	0				
Gly	6	3	0				
Met	4	6	0				
His	5	3	0				
	166	166	48	24	18	5	1

*The data are from Taniguchi et al. (ref. 21) and Mantei et al. (ref. 19).

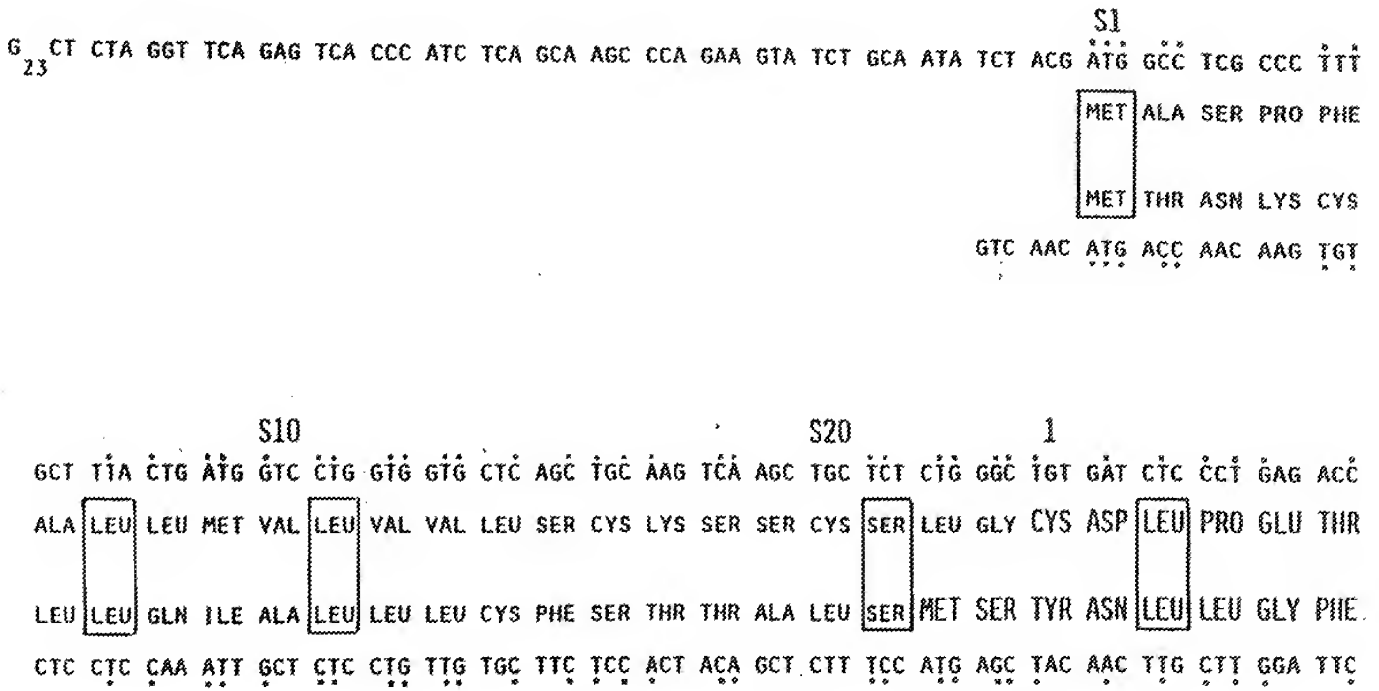
FIGURE LEGENDS

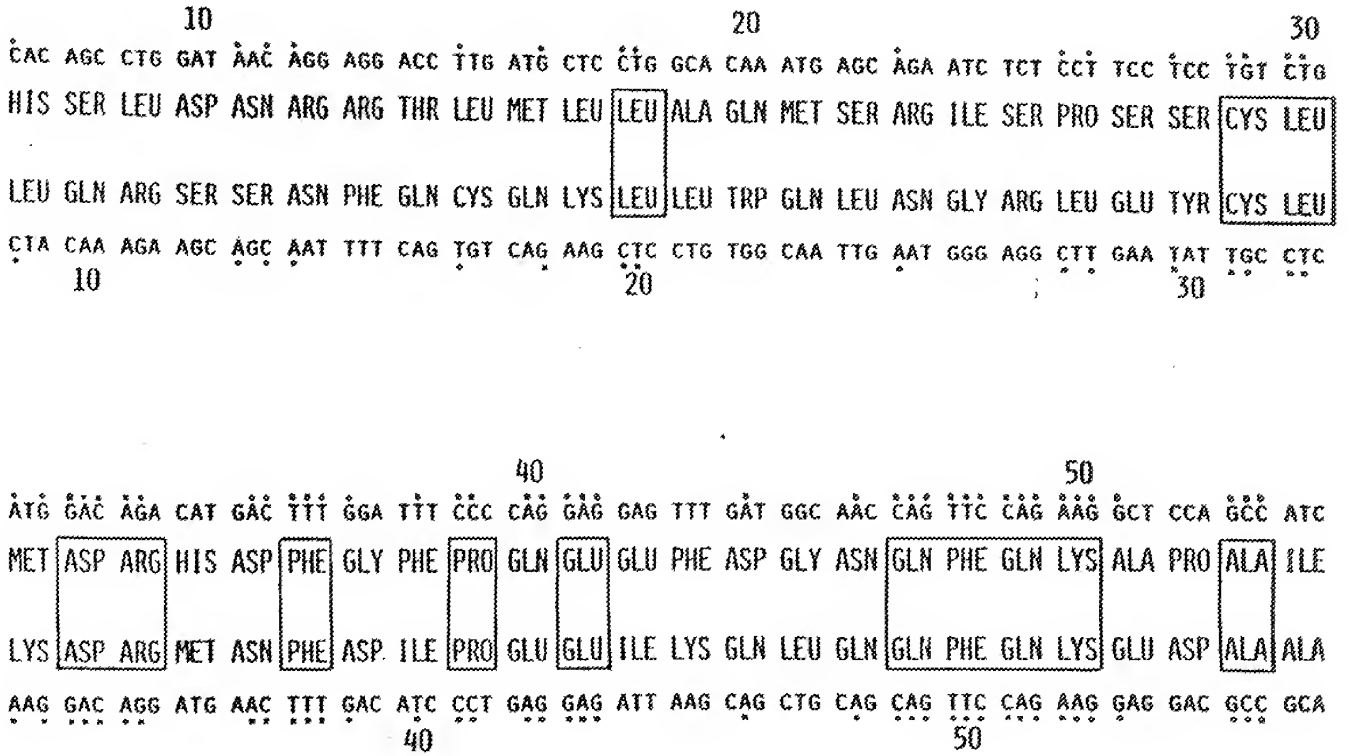
Fig. 1 Comparison of the nucleotide sequences of human leukocyte interferon I (Le-IF I) and human fibroblast interferon cDNA and of the derived amino acid sequences. The sequences are from Mantei et al.(19) and Taniguchi et al. (21). They were aligned to give maximal homology. Identical amino acids are framed, identical nucleotides are marked by a dot. S1 to S23 indicate the amino acids of the putative signal sequence; 1 to 166 the amino acids of the interferon polypeptides.

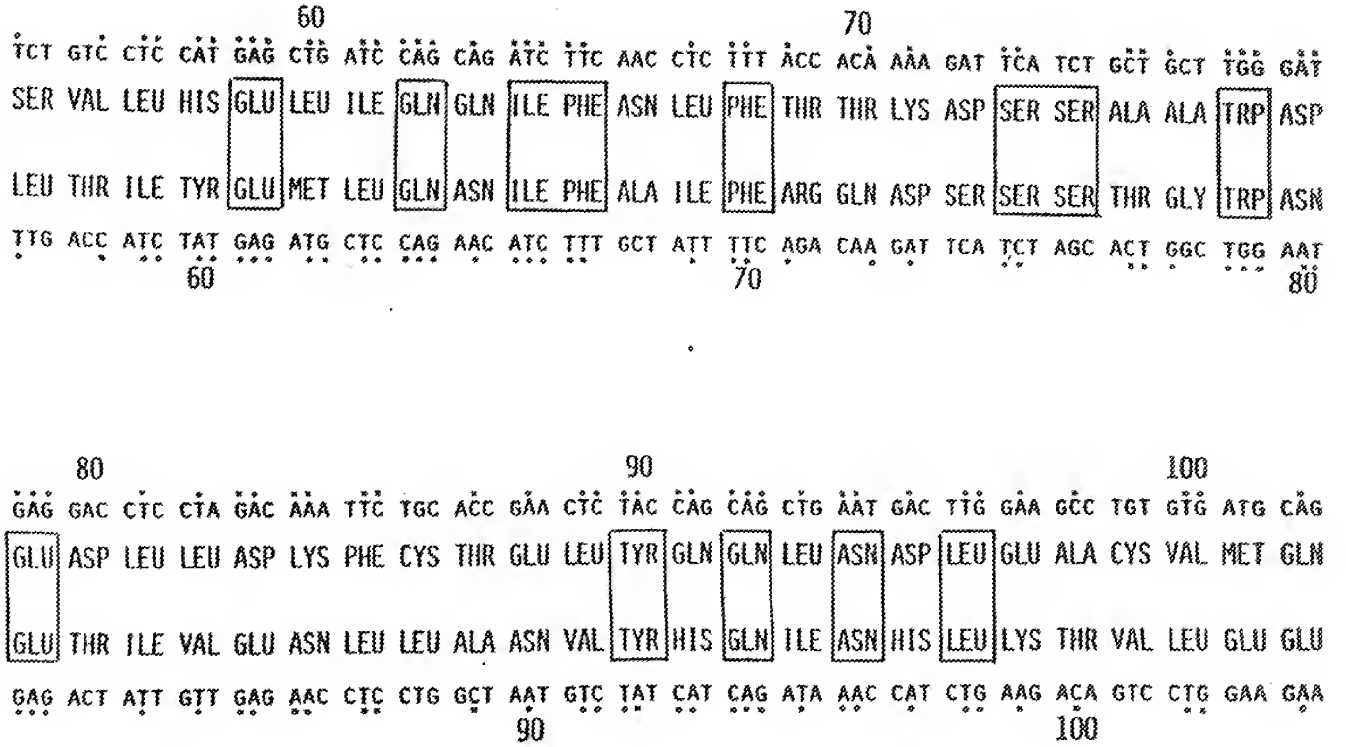
Fig. 2 Similarity of the nucleotide and amino acid sequences of human leukocyte interferon I and fibroblast interferon.

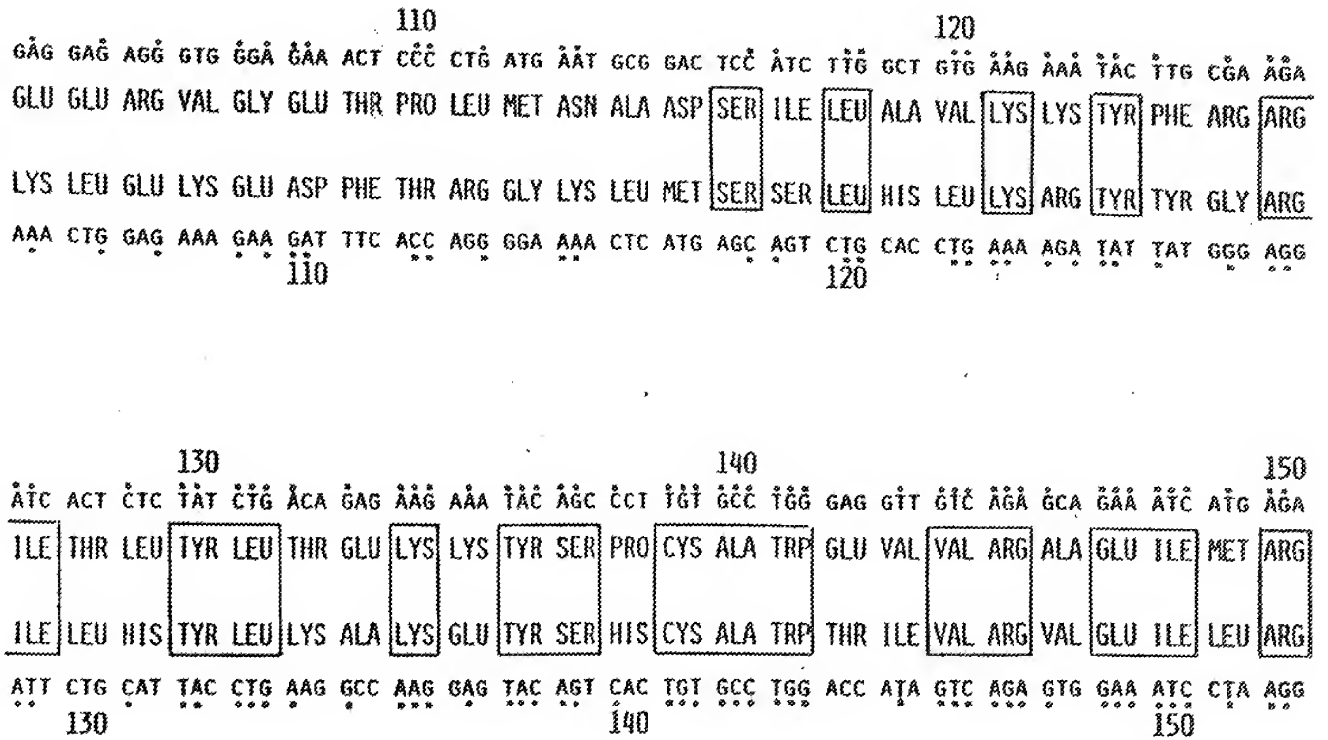
The sequences shown in Fig. 1 were subdivided in segments of 8 amino acids or 24 nucleotides, each overlapping by 4 and 12 residues, respectively, with the neighboring segments. The percentage of coincident residues was plotted as a function of map position. Open vertical blocks, nucleotides; filled vertical blocks, amino acids. L-IF, leukocyte interferon cDNA; F-IF, fibroblast interferon cDNA; lines, non-coding sequences; hatched bars, putative signal peptide; open bars, interferon polypeptide.

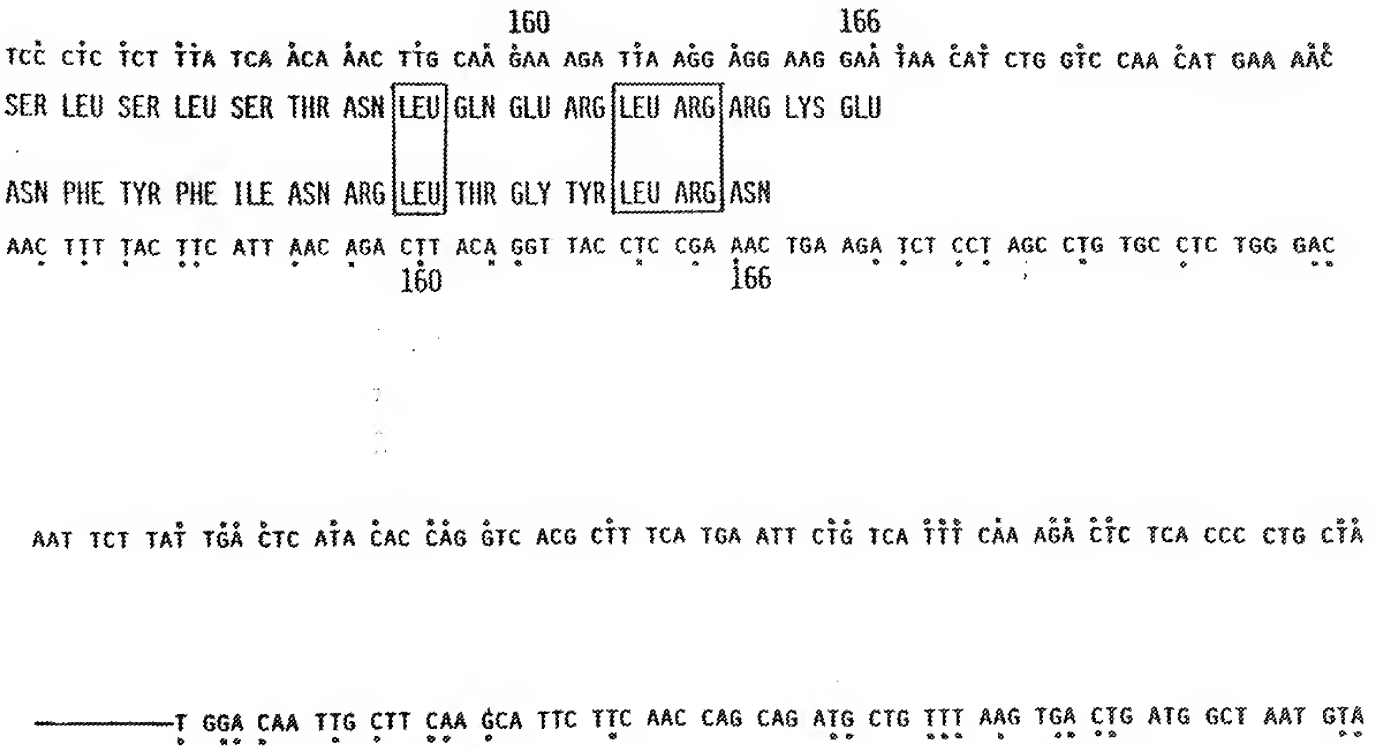
Fig. 1


S1
G23CT CTA GGT TCA GAG TCA CCC ATC TCA GCA AGC CCA GAA GTA TCT GCA ATA TCT ACG ATG GCC TCG CCC TTT
MET ALA SER PRO PHE
MET THR ASN LYS CYS
GTC AAC ATG ACC AAC AAG TGT
S10
S20
1
GCT TTA CTG ATG GTC CTG GTG GTG CTC AGC TGC AAG TCA AGC TGC TCT CTG GGC TGT GAT CTC CCT GAG ACC
ALA LEU LEU MET VAL LEU VAL VAL LEU SER CYS LYS SER SER CYS SER LEU GLY CYS ASP LEU PRO GLU THR
LEU LEU GLN ILE ALA LEU LEU LEU CYS PHE SER THR THR ALA LEU SER MET SER TYR ASN LEU LEU GLY PHE
CTC CTC CAA ATT GCT CTC CTG TTG TGC TTC TCC ACT ACA GCT CTT TCC ATG AGC TAC AAC TTG CTT GGA TTC
1


10
20
30
CAC AGC CTG GAT AAC AGG AGG ACC TTG ATG CTC CTG GCA CAA ATG AGC AGA ATC TCT CCT TCC TCC TGT CTG
HIS SER LEU ASP ASN ARG ARG THR LEU MET LEU LEU ALA GLN MET SER ARG ILE SER PRO SER SER CYS LEU
LEU GLN ARG SER SER ASN PHE GLN CYS GLN LYS LEU LEU TRP GLN LEU ASN GLY ARG LEU GLU TYR CYS LEU
CTA CAA AGA AGC AGC AAT TTT CAG TGT CAG AAG CTC CTG TGG CAA TTG AAT GGG AGG CTT GAA TAT TGC CTC
10
20
30
40
50
ATG GAC AGA CAT GAC TTT GGA TTT CCC CAG GAG GAG TTT GAT GGC AAC CAG TTC CAG AAG GCT CCA GCC ATC
MET ASP ARG HIS ASP PHE GLY PHE PRO GLN GLU GLU PHE ASP GLY ASN GLN PHE GLN LYS ALA PRO ALA ILE
LYS ASP ARG MET ASN PHE ASP ILE PRO GLU GLU ILE LYS GLN LEU GLN GLN PHE GLN LYS GLU ASP ALA ALA
AAG GAC AGG ATG AAC TTT GAC ATC CCT GAG GAG ATT AAG CAG CTG CAG CAG TTC CAG AAG GAG GAC GCC GCA
40
50


60
70
TCT GTC CTC CAT GAG CTG ATC CAG CAG ATC TTC AAC CTC TTT ACC ACA AAA GAT TCA TCT GCT GCT TGG GAT
SER VAL LEU HIS GLU LEU ILE GLN GLN ILE PHE ASN LEU PHE THR THR LYS ASP SER SER ALA ALA TRP ASP
LEU THR ILE TYR GLU MET LEU GLN ASN ILE PHE ALA ILE PHE ARG GLN ASP SER SER SER THR GLY TRP ASN
TTG ACC ATC TAT GAG ATG CTC CAG AAC ATC TTT GCT ATT TTC AGA CAA GAT TCA TCT AGC ACT GGC TGG AAT
60
70
80
80
90
100
GAG GAC CTC CTA GAC AAA TTC TGC ACC GAA CTC TAC CAG CAG CTG AAT GAC TTG GAA GCC TGT GTG ATG CAG
GLU ASP LEU LEU ASP LYS PHE CYS THR GLU LEU TYR GLN GLN LEU ASN ASP LEU GLU ALA CYS VAL MET GLN
GLU THR ILE VAL GLU ASN LEU LEU ALA ASN VAL TYR HIS GLN ILE ASN HIS LEU LYS THR VAL LEU GLU GLU
GAG ACT ATT GTT GAG AAC CTC CTG GCT AAT GTC TAT CAT CAG ATA AAC CAT CTG AAG ACA GTC CTG GAA GAA
90
100


110
120
GAG GAG AGG GTG GGA GAA ACT CCC CTG ATG AAT GCG GAC TCC ATC TTG GCT GTG AAG AAA TAC TTG CGA AGA
GLU GLU ARG VAL GLY GLU THR PRO LEU MET ASN ALA ASP SER ILE LEU ALA VAL LYS LYS TYR PHE ARG ARG
LYS LEU GLU LYS GLU ASP PHE THR ARG GLY LYS LEU MET SER SER LEU HIS LEU LYS ARG TYR TYR GLY ARG
AAA CTG GAG AAA GAA GAT TTC ACC AGG GGA AAA CTC ATG AGC AGT CTG CAC CTG AAA AGA TAT TAT GGG AGG
110
120
130
140
150
ATC ACT CTC TAT CTG ACA GAG AAG AAA TAC AGC CCT TGT GCC TGG GAG GTT GTC AGA GCA GAA ATC ATG AGA
ILE THR LEU TYR LEU THR GLU LYS LYS TYR SER PRO CYS ALA TRP GLU VAL VAL ARG ALA GLU ILE MET ARG
ILE LEU HIS TYR LEU LYS ALA LYS GLU TYR SER HIS CYS ALA TRP THR ILE VAL ARG VAL GLU ILE LEU ARG
ATT CTG CAT TAC CTG AAG GCC AAG GAG TAC AGT CAC TGT GCC TGG ACC ATA GTC AGA GTG GAA ATC CTA AGG
130
140
150


160
166
TCC CTC TCT TTA TCA ACA AAC TTG CAA GAA AGA TTA AGG AGG AAG GAA TAA CAT CTG GTC CAA CAT GAA AAC
SER LEU SER LEU SER THR ASN LEU GLN GLU ARG LEU ARG ARG LYS GLU
ASN PHE TYR PHE ILE ASN ARG LEU THR GLY TYR LEU ARG ASN
AAC TTT TAC TTC ATT AAC AGA CTT ACA GGT TAC CTC CGA AAC TGA AGA TCT CCT AGC CTG TGC CTC TGG GAC
160
166
AAT TCT TAT TGA CTC ATA CAC CAG GTC ACG CTT TCA TGA ATT CTG TCA TTT CAA AGA CTC TCA CCC CTG CTA
T GGA CAA TTG CTT CAA GCA TTC TTC AAC CAG CAG ATG CTG TTT AAG TGA CTG ATG GCT AAT GTA

-TA ACT ATG ACC ATG CTG ATA AAC TGA TTT ATC TAT TTA AAT ATT TAT TTA ACT ATT CAT AAG ATT TAA ATT

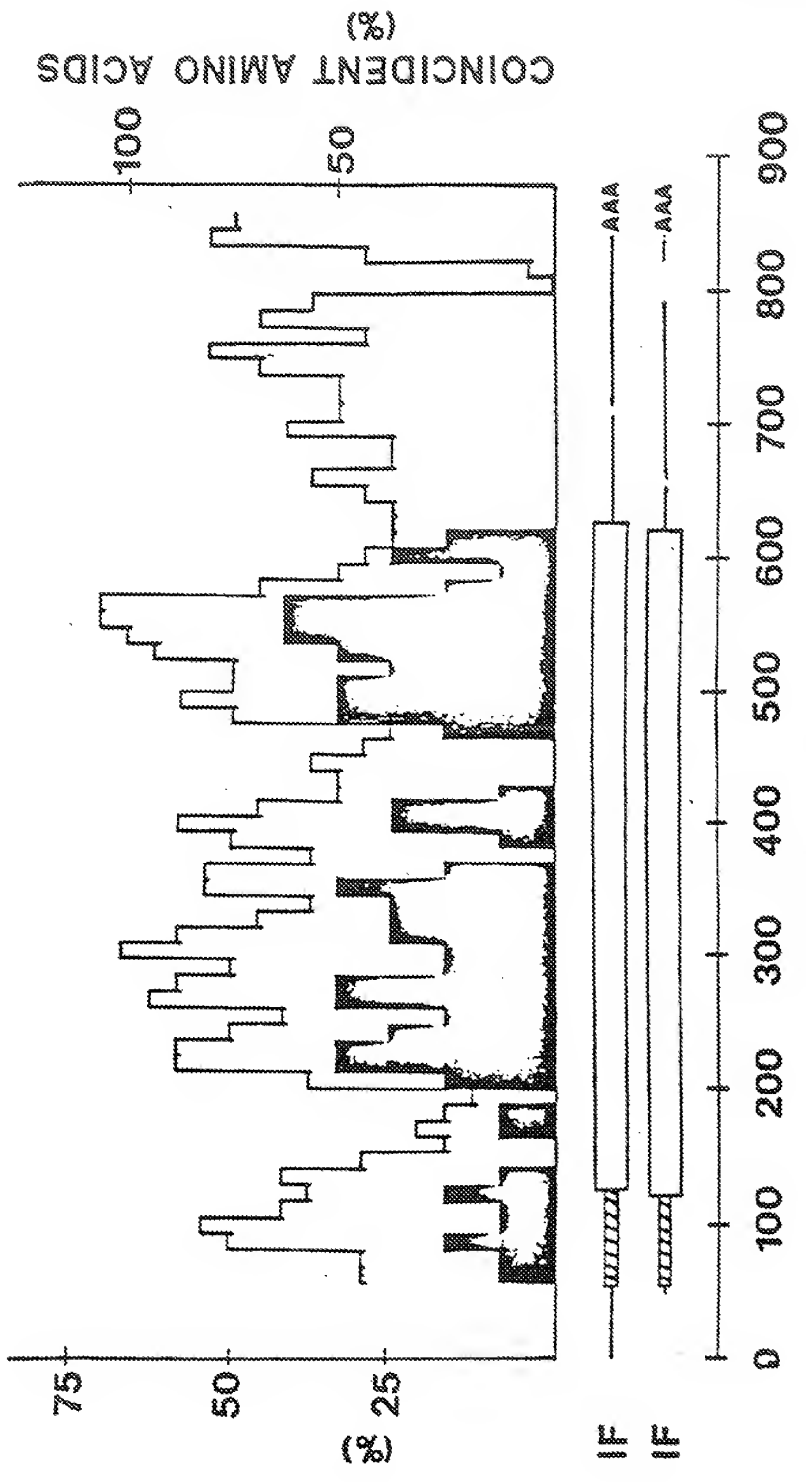
CTG CAT ATG AAA GGA CAC TAG AAG ATT TTG AAA TTT TTA TTA AAT TAT GAG TTA TTT TTA TTT ATT TAA ATT

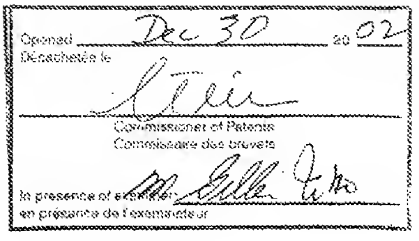
ATT TTT GTT CAT ATA ACG TCA TGT GCA CCT TTA CAC TGT GGT TAG TGT AAT AAA ACA TGT TCC TTA TAT TTA

---TT ATT TTG GAA AAT AAA TTA TTT TTG GTG CAA A-A

CTC AAA AAA A

GTC AAA A


COINCIDENT AMINO ACIDS
(%)
100
50
75
50
25
(%)
IF
IF
AAA
AAA
0
100
200
300
400
500
600
700
800
900
BASE PAIRS



This is EXHIBIT FIERS-41
to
the Affidavit of Walter C. Fiers
sworn before me
this 15 th day of November, 2001

Commissioner for Oath or Notary Public

40 [illegible]
41 Maxam, A. M. & Gilbert, W. Proc. natn. Acad. Sci. U.S.A. 74, 560–564 (1977).
42 Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. Nature 270, 486–494 (1977).
43 Proudfoot, N. J. & Brownlee, G. G. Nature 263, 211–214 (1976).
44 Fiers, W. et al. Nature 273, 113–120 (1978).
45 [illegible] Nature 283, [illegible] (1980).
46 Min Jou, W. et al. Cell 19, 683–696 (1980).
47 Shine, J., Seeburg, P. H., Martial, J. A., Baxter, J. D. & Goodman, H. M. Nature 270, 494–499 (1977).
48 [illegible]
49 [illegible] Proc. natn. Acad. Sci. U.S.A. 75, [illegible]
50 [illegible] J. biol. Chem. 251, [illegible]
51 Neuberger, A. [illegible]
52 [illegible] Proc. natn. Acad. Sci. U.S.A. 72, [illegible]
53 Taniguchi, T. et al. Gene 10, 11–15 (1980).
54 Taniguchi, T. et al. Nature 285, 547–549 (1980).
55 Mantei, N. et al. Gene 10, 1–10 (1980).
56 [illegible]

Human leukocyte and fibroblast interferons are structurally related

Tadatsugu Taniguchi*, Ned Mantei†, Marco Schwarzstein†, Shigekazu Nagata†, Masami Muramatsu* & Charles Weissmann†

*Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Tokyo 170, Japan

†Institut für Molekularbiologie I, Universität Zürich, 8093 Zürich, Switzerland

The coding sequences of the cDNAs of cloned human leukocyte interferon I and human fibroblast interferon show homologies of 45% at the nucleotide and 29% at the amino acid level. We conclude that the two genes were derived from a common ancestor.

THE acid-stable human interferons are subdivided into two major groups, fibroblast interferons (F-IF) and leukocyte interferons (Le-IF); these are the major components of the interferons produced by induced fibroblasts and leukocytes, respectively. Some cells, such as the lymphoblastoid Namalva cell line, produce a mixture of 90% Le-IF and 10% F-IF[1,2]. The two interferon types have several features in common: both are glycoproteins with molecular weights ranging from 16,000 to 26,000 (refs 3–9), the induction and shut-off of their synthesis seem to be under similar control[6], and at least some of the responses elicited in target cells are similar, such as induction of an antiviral state which is accompanied by increased synthesis of several proteins[10–13]. Nonetheless, the two kinds of interferon differ in many respects. Antibodies directed against Le-IF do not neutralize F-IF and vice versa[14], the target cell specificities of the two interferons differ[15] and the sequences of the 13 amino-terminal amino acids of F-IF and Le-IF (from lymphoblastoid cells) show no homology[16,17]. Although Le-IF and F-IF are encoded by different mRNA species[18], it is not known whether these mRNAs are transcribed from distinct genes or whether they arise from the same gene through a common precursor which is processed or spliced in different modes.

We have recently cloned and sequenced one species each of Le-IF (Le-IF I)[19,20] and F-IF cDNA[21,22]. A second species of Le-IF (Le-IF II) cDNA has recently been identified (M. Streuli, S.N. and C.W., unpublished results).

Comparison of the amino acid sequences of Le-IF and F-IF

In Fig. 1 the nucleotide sequences of Le-IF I and F-IF cDNA were aligned so that the AUGs closest to their 5' termini coincided. From the amino-terminal sequence published for F-IF[16] and lymphoblastoid Le-IF[17], one can determine that for F-IF and Le-IF, respectively, the 21st and 23rd codons following the initiation triplet represent the first amino acid of the interferon polypeptide. Presumably, the stretch in between encodes a signal peptide. As the respective putative signal peptides of Le-IF and F-IF comprise 23 and 21 amino acids, the IF polypeptides, as aligned in Fig. 1, are shifted by two residues relative to their termini. In this alignment, 48 of 166 positions (29%) of the interferon polypeptides have identical amino acids. By introducing appropriate gaps, better homology could be achieved, particularly in the region of the signal sequence; in the present comparison this has not been done.

To plot the degree of homology between the F-IF and Le-IF as a function of the map distance, the sequence was subdivided into segments of 8 amino acids (or 24 nucleotides), each overlapping by 4 amino acids (or 12 nucleotides) with the neighbouring segments, and the per cent coincidence of amino acids (and nucleotides) for each segment was determined (see ref. 23). As seen in Fig. 2, amino acid sequences show three domains of homology. The first one, with the least degree of homology, corresponds to the putative signal sequence, which is rich in hydrophobic residues and has 4 identical amino acid positions out of 21; the second domain, between amino acids 28 and 80 (counted on the Le-IF sequence), has 21 identical residues out of 51 (41% homology); and the third, between positions 115 and 151 (Le-IF sequence), has 19 out of 35 identical residues (54%). The longest stretches of contiguous conserved amino acids are Gln-Phe-Gln-Lys (positions 47–50 of Le-IF and 49–52 of F-IF) and Cys-Ala-Trp (positions 139–141 and 141–143, respectively); the latter sequence is notable because it comprises Cys and Trp, which are preferentially conserved in related proteins[24]. Table 1 shows that conservation was highest between the interferon polypeptides (not considering the signal sequences) for Trp, Phe, Arg, Cys and Tyr residues, in agreement with the general experience that the amino acids most likely to be conserved between related proteins are Trp > Cys > Tyr > Arg > Phe, His (ref. 24). Even where amino acids are conserved, the codons show one or more nucleotide changes in half the instances. The codons of three out of seven conserved Leu residues are unrelated, as are two of four codons pertaining to conserved Ser residues. This suggests that there is a strong selective pressure favouring the conservation of several amino acids. It is quite likely that at least some of the conserved amino acids are essential for a function common to Le-IF and F-IF, perhaps the induction of the virus-resistant state in the target cell. These findings may provide guidelines for the tailoring of modified[25], possibly shorter polypeptides possessing certain activities of interferon.

Comparison of the nucleotide sequences of Le-IF and F-IF

The nucleic acid sequences show an average homology of 43% in the domain of the signal sequence and of 45% in the interferon polypeptide sequence. On a random basis, about

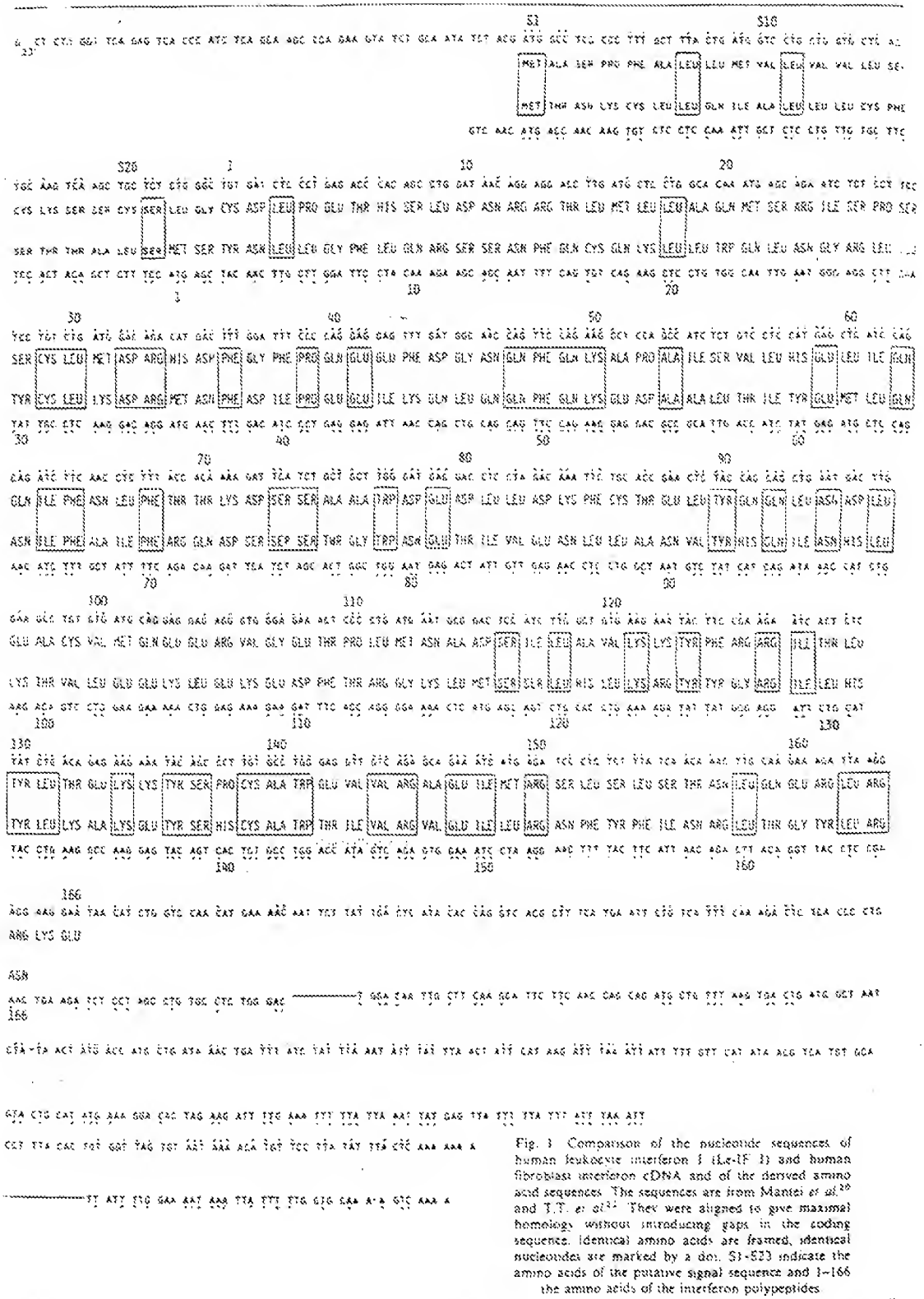


Fig. 1 Comparison of the nucleotide sequences of human leukocyte interferon 1 (Le-IF 1) and human fibroblast interferon cDNA and of the derived amino acid sequences. The sequences are from Mantei *et al.*[20] and T.T. *et al.*[21]. They were aligned to give maximal homology without introducing gaps in the coding sequence. Identical amino acids are framed, identical nucleotides are marked by a dot. S1–S23 indicate the amino acids of the putative signal sequence and 1–166 the amino acids of the interferon polypeptides.

25%, of the nucleotide positions should coincide. Within the interferon coding sequence the nucleotide homologies are more evenly distributed than the amino acid homologies. However, one may distinguish, albeit to a less pronounced degree, the same three blocks of similarity noted for the amino acids. The longest region without mismatches extends for 33 nucleotides (compare 47th to 51st codon of Le-IF with 49th to 53rd codon of F-IF). There are, in addition, sequences of 17, 18 and 20 nucleotides with 1, 3 and 4 mismatches, respectively. The heteropolymeric 3'-terminal noncoding region of Le-IF cDNA has 342 nucleotides, and is longer by 39 residues than its counterpart in F-IF cDNA. In aligning the two sequences, four gaps were introduced to maximize homology, as described by van Ooyen *et al.*[23]. In this way, several segments were matched with 29–41% homology. The introduction of gaps in the alignment may be justified in view of the arguments presented previously, that intervening sequences and noncoding regions of reduplicated genes diverge as a consequence of block insertions and/or deletions in the course of evolution[23,24]. It is unlikely that the extent of homology between Le-IF and F-IF cDNA would allow significant cross-hybridization between the two species.

A common ancestral gene for Le-IF and F-IF

On the basis of our findings there is no doubt that Le-IF and F-IF genes are derived from a common ancestral sequence. When did the separation of these genes occur? Human α- and β-globin show 57% amino acid mismatches, and human β-globin and myoglobin, as well as α-globin and myoglobin, 86% mismatches. If the rate of divergence of interferons and globins is comparable (however, this is quite uncertain, see ref. 24 p. 50 for proteins showing both higher and lower rates), the separation of interferon genes occurred after that of myoglobin and haemoglobins but before that of α- and β-globins, that is between 500 and 1,000 Myr ago[24], which is about the time vertebrates arose[25]. This would mean that both types of interferon gene should occur in all vertebrates, unless one and/or the other was lost by deletion. Indeed, as shown by the sequencing of 17–24 amino-terminal residues[9,11,26], mouse interferons A and B show significant homology with human fibroblast interferon, and mouse interferon C with human lymphoblastoid interferon, whereas the mouse species A and B on the one hand, and the species C on the other show no significant homology within the short segment sequenced. Thus, at least in the mouse, representatives of both interferon families exist. It will be of interest to determine the evolutionary relationship of these to the third type of interferon, immune or γ-interferon.

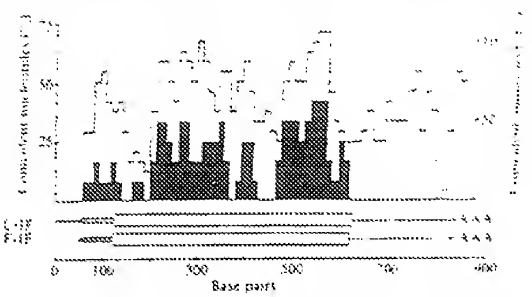


Fig. 2 Similarity of the nucleotide and amino acid sequences of human leukocyte interferon 1 and fibroblast interferon. The sequences shown in Fig. 1 were subdivided into segments of 5 amino acids or 20 nucleotides, each overlapping by 4 and 12 residues, respectively, with the neighbouring segment. The percentage of coincident residues was plotted as a function of map position. Open vertical blocks, nucleotides; filled vertical blocks, amino acids. L-IF, leukocyte interferon cDNA; F-IF, fibroblast interferon cDNA. lines, noncoding sequences; hatched bars, putative signal peptide; open bars, interferon polypeptide.

After submission of this article, we learnt that Derynck *et al.* had cloned and sequenced fibroblast interferon (see accompanying article[27]), confirming the deduced amino acid sequence of T.T *et al.*[22].

T.T. and M.M. thank Dr H. Sugano for his continuous support and interest. The work at the University of Zurich was supported by Biogen SA and the Schweizerische Nationalfonds.

Table 1 Conservation of amino acids in leukocyte and fibroblast interferon

	F-IF	Le-IF	Conserved amino acids	No. of changes in codon of conserved amino acids 0	1	2	3
Leu	25	22	8	1	4	3	
Cys	3	3	2	1	1		
Asn	12	6	1	1			
Arg	13	12	5	1	3	1	
Phe	9	8	4	2	2		
Pro	7	6	1		1		
Gln	11	10	3	3			
Lys	11	8	3	2	1		
Asp	9	10	2	2			
Glu	17	15	6	6			
Ile	11	7	2	2	5		
Ser	9	13	4		2	1	1
Trp	3	2	2	2			
Tyr	10	4	4	1	3		
Val	5	8	3	1			
Asp	5	11	3	1			
Thr	6	9	0				
Gly	6	7	0				
Met	4	6	0				
His	3	5	0				
Total	166	166	48	24	18	5	1

The data are from T.T. *et al.*[22] and Mantei *et al.*[19].

Received 25 March; accepted 29 April 1980.

1. Havell, E. A., Vilcek, J. & Gifford, J. *J. gen. Virol.* **38**, 53–59 (1978).
2. Paucker, K. *Yale Res. Bull. Med.* **28**, 23–26 (1972).
3. Brosius, R. S. *et al. J. biol. Chem.* **252**, 6585–6588 (1977).
4. Knight, E. Jr *Proc. natn. Acad. Sci. U.S.A.* **73**, 520–523 (1976).
5. Lee, S. H. *et al. J. biol. Chem.* **254**, 8067–8073 (1979).
6. Havell, E. G., Vilcek, J. & Vilcek, J. *Virology* **68**, 383–393 (1976).
7. Rubinstein, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, in the press (1979).
8. Stewart, W. E. II & Wheelock-Stewart, M. *Virology* **97**, 473–478 (1979).
9. Havell, E. A., Sarmiento, S. & Gifford, J. *J. biol. Chem.* **252**, 4425–4427 (1977).
10. Lebleu, B. Sen, G. C., Shaila, S. Cabrer, B. & Lengyel, P. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3107–3111 (1976).
11. Shepardson, A. D., Borden, E. F. & Kerr, I. M. *Science* **205**, 1375–1377 (1979).
12. Hovanessian, A. G. & Kerr, I. M. *Eur. J. Biochem.* **93**, 515–526 (1979).
13. Schmidt, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 4788–4792 (1979).
14. Havell, E. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 2185–2188 (1975).
15. Stewart, W. E. II in *The Interferon System*, 134–143 (Springer, New York, 1979).
16. Knight, E. Jr, Hunkapiller, M. W., Korant, B. D., Hardy, R. W. F. & Hood, L. E. *Science* **207**, 525–526 (1980).
17. Zoon, K. C. *et al. Science* **207**, 527–528 (1980).
18. Cavalieri, R. L., Havell, E. A., Vilcek, J. & Pestka, S. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3287–3291 (1977).
19. Nagata, S. *et al. Nature* **284**, 316–320 (1980).
20. Mantei, N. *et al. Gene* **10**, 1–10 (1980).
21. Taniguchi, T. *et al. Proc. Jpn Acad.* **55B**, 464–469 (1979).
22. Taniguchi, T., Ohno, S., Fujii-Kuriyama, Y. & Muramatsu, M. *Gene* **10**, 11–15 (1980).
23. van Ooyen, A., van den Berg, J., Mantei, N. & Weissmann, C. *Science* **206**, 337–344 (1979).
24. Dayhoff, M. O. in *Atlas of Protein Sequence and Structure* Vol. 5 (Natn. Biomed. Res. Fdn, Washington, 1972).
25. Müller, W., Weber, H., Meyer, F. & Weissmann, C. *J. molec. Biol.* **124**, 343–358 (1978).
26. Knorre, D. A., Manuilov, I. V. & Lesser, P. *Cell* **18**, 883–893 (1979).
27. Kiefhaber, B. in *Proc. XXII int. Congr. Zool.* Vol. 6 (ed. Moore, J. A.) 327–334 (Natural History Press, New York, 1965).
28. Taira, H. *et al. Science* **205**, 528–530 (1980).
29. Derynck, R. *et al. Nature* **285**, 542–547 (1980).

SOUTH AUSTRALIAN INSTITUTE
OF TECHNOLOGY

TERRACE, ADELAIDE, SOUTH AUSTRALIA . . . . . TELEPHONE XXX XXXX

School of Pharmacy

10th April, 1980

Dr. C. Weissmann,
Institut fur Molecularbiologie I,
Universitat Zurich,
HONGGERBERG. Zurich. 8093
Switzerland.

18 April 1980

Dear Dr. Weissmann,

Thank you again for the preprints. Fantastic work. Have you thought about the possible advantages of interferon production in B. subtilis? The leader peptides may be adequate for transport of the interferons out to the medium thus simplifying the purification of interferon (proteases may be a problem). A sequence comparison realignment of sections of the interferons is enclosed. I was quite surprised by the similarity in amino acid composition of bovine growth hormone and Le-ifl. However I could find little sequence homology.

I wish you all the very best in your work.

Yours sincerely,

Roy

R.J. Harris
Lecturer in Biochemistry

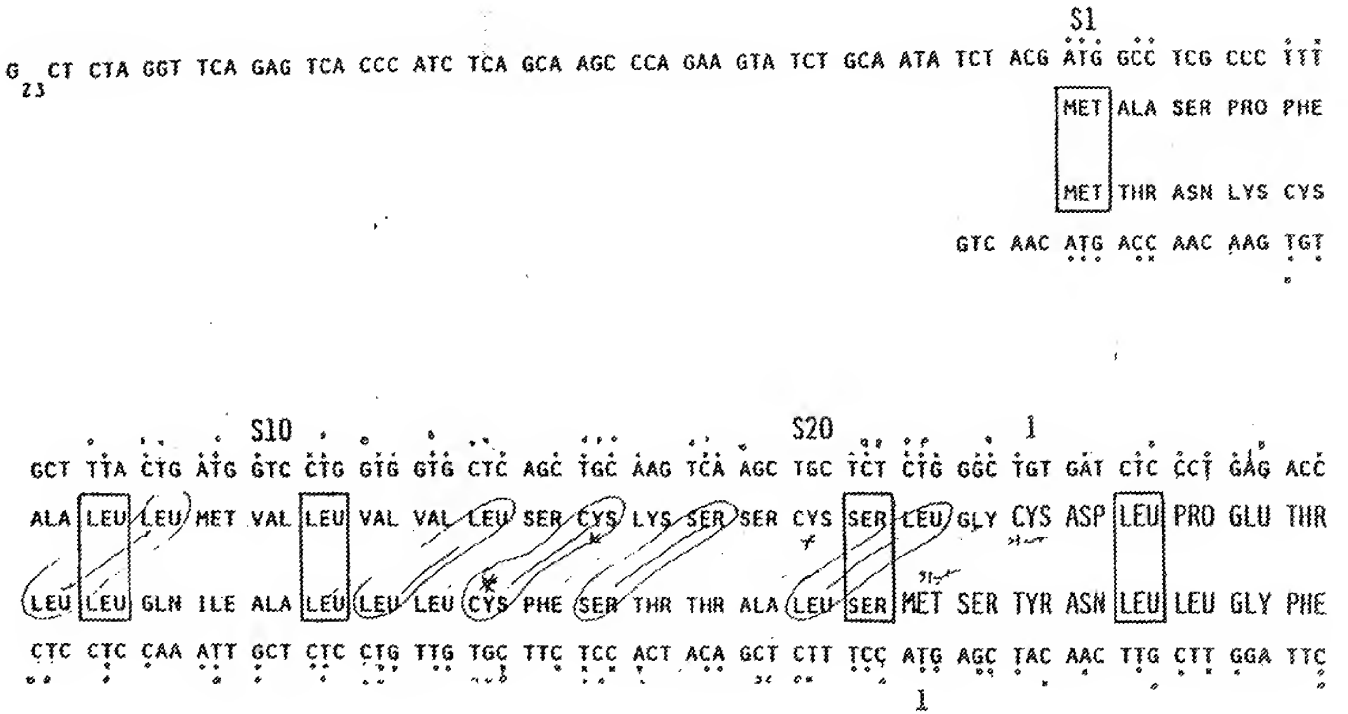
PS. Do you mind if I distribute the preprints within Australia? I will assume this is ok unless you notify me otherwise. I will be presenting data on applications of genetic engineering to a meeting in May 1980. I would love to work in your lab but this would perhaps be incompatible with my aim to produce interferon within Australia.
Would you be interested in attending the International Biochem. Conference in Perth (Western Australia) 1982?

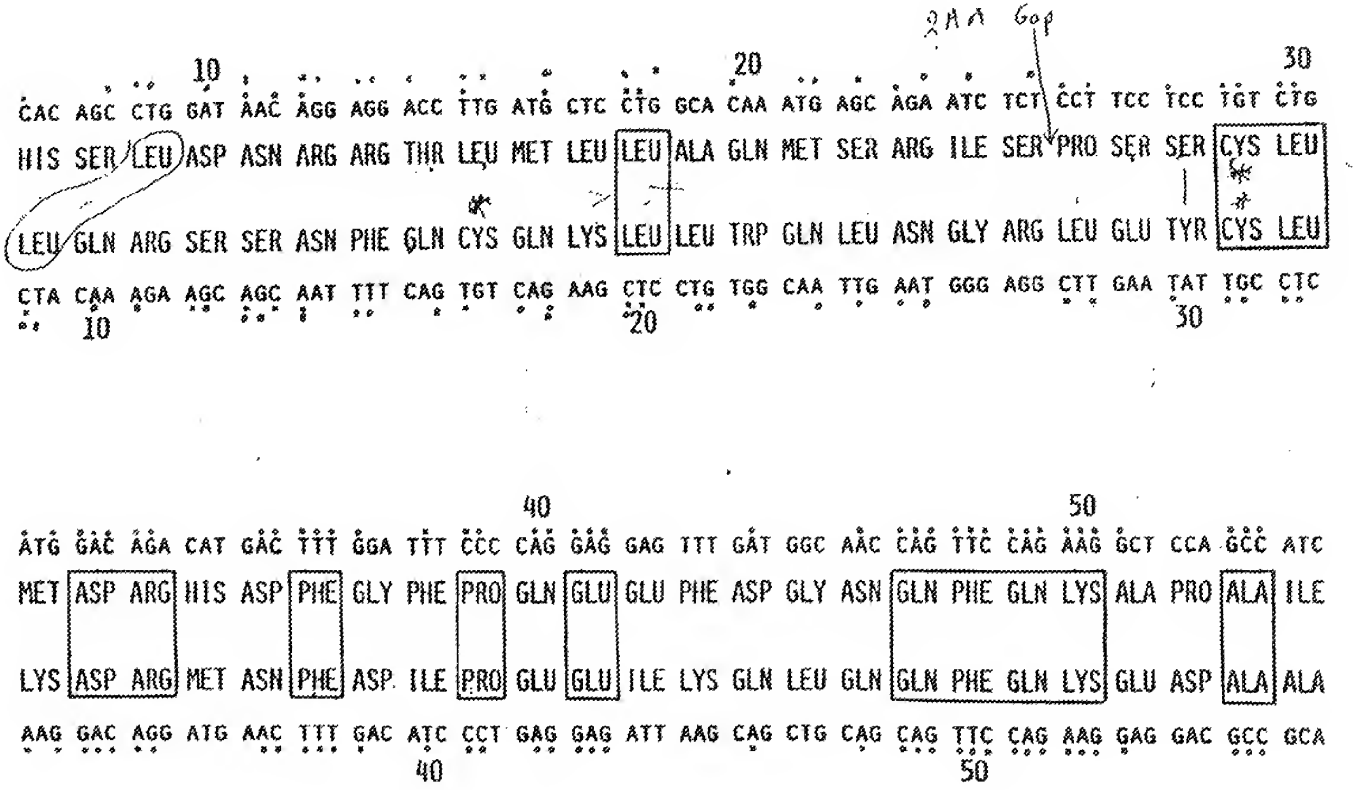
This is EXHIBIT FIERS-42
to
the Affidavit of Walter C. Fiers
sworn before me
this 19 th day of November, 2001

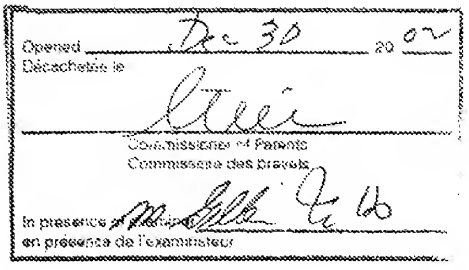
Commissioner for Oath or Notary Public

Fig. 1

ds complete run




10
20
30
CAC AGC CTG GAT AAC AGG AGG ACC TTG ATG CTC CTG GCA CAA ATG AGC AGA ATC TCT CCT TCC TCC TGT CTG
HIS SER LEU ASP ASN ARG ARG THR LEU MET LEU LEU ALA GLN MET SER ARG ILE SER PRO SER SER CYS LEU
LEU GLN ARG SER SER ASN PHE GLN CYS GLN LYS LEU LEU TRP GLN LEU ASN GLY ARG LEU GLU TYR CYS LEU
CTA CAA AGA AGC AGC AAT TTT CAG TGT CAG AAG CTC CTG TGG CAA TTG AAT GGG AGG CTT GAA TAT TGC CTC
10
20
30
40
50
ATG GAC AGA CAT GAC TTT GGA TTT CCC CAG GAG GAG TTT GAT GGC AAC CAG TTC CAG AAG GCT CCA GCC ATC
MET ASP ARG HIS ASP PHE GLY PHE PRO GLN GLU GLU PHE ASP GLY ASN GLN PHE GLN LYS ALA PRO ALA ILE
LYS ASP ARG MET ASN PHE ASP ILE PRO GLU GLU ILE LYS GLN LEU GLN GLN PHE GLN LYS GLU ASP ALA ALA
AAG GAC AGG ATG AAC TTT GAC ATC CCT GAG GAG ATT AAG CAG CTG CAG CAG TTC CAG AAG GAG GAC GCC GCA
40
50



This is EXHIBIT FIERS-43
to
the Affidavit of Walter C. Fiers
sworn before me
this 13 th day of November, 2001

Commissioner for Oath or Notary Public

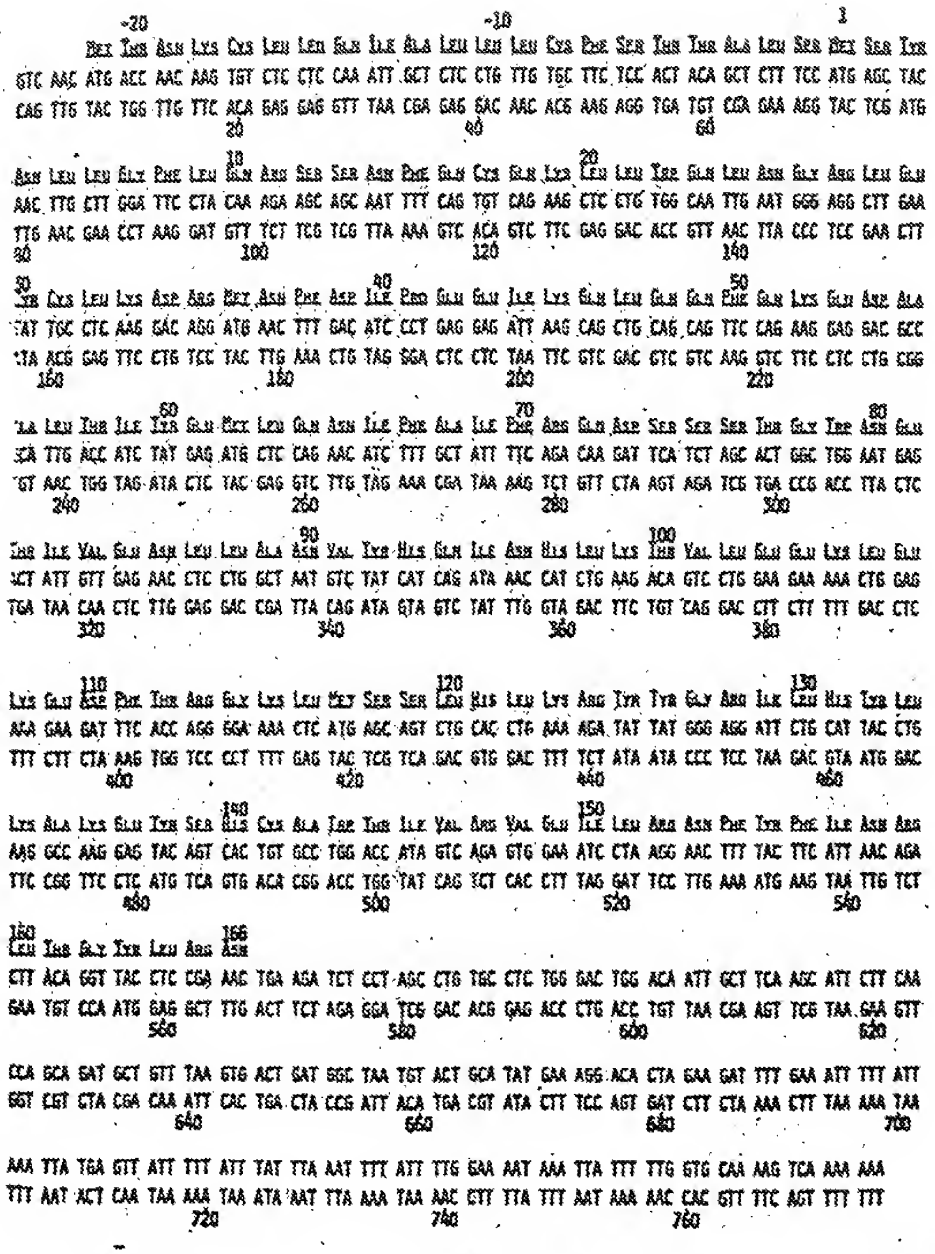


Fig. 2. Nucleotide sequence of the human fibroblast interferon cDNA and the predicted amino acid sequence of the pre-interferon molecule. Determination of the DNA sequence was carried out according to the procedure of Maxam and Gilbert (1977). (cf. Fig. 1). The DNA sequence permits prediction of the entire amino acid sequence for human fibroblast interferon (amino acids 1–166) and its putative signal peptide (amino acids –21 to –1) as shown in the line above the DNA sequences.

mann and A. Fuse for valuable discussions and advice. Thanks are also due to Ms. Y. Taniguchi for typing the manuscript.

REFERENCES

Berg, K., Ogburn, C.A., Paucker, K., Mogensen, K.E. and Cantell, K.: Affinity chromatography of human leukocyte and diploid cell interferons on sepharose-bound antibodies. J. Immunol. 114 (1975) 640–644.

Cavalieri, R.L., Havell, E.A., Vilček, J. and Pestka, S.: Synthesis of human interferon by *Xenopus laevis* oocytes: two structural genes for interferons in human cells. Proc. Natl. Acad. Sci. USA 74 (1977) 3287–3291.

Currier, T.C. and Nester, E.W.: Isolation of covalently closed circular DNA of high molecular weight from bacteria, Analyt. Biochem. 76 (1976) 431–441.

De Maeyer, E. and De Maeyer-Guignard, J.: Effect of interferon on cell-mediated immunity. Texas Repts. Biol. Med. 35 (1977) 370–374.

Epstein, L.B.: Mitogen and antigen induction of interferon in vitro and in vivo. Texas Repts. Biol. Med. 35 (1977) 42–56.

Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voorde, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G. and Ysebaert, M.: Complete nucleotide sequence of SV40 DNA. Nature 273 (1978) 113–120.

Grantham, R., Gautier, C., Gouy, M., Mercier, R. and Pavé, A.: Codon catalog usage and the genome hypothesis. Nucl. Acids Res. 8 (1980) r49–r62.

Gresser, I., Bandu, M.T., Brouty-Boye, D. and Tovey, M.: Pronounced antiviral activity of human interferon on bovine and porcine cells. Nature 251 (1974) 543–545.

Gresser, I.: On the mechanism of the antitumor effects of interferon. Texas Repts. Biol. Med. 35 (1977) 394–405.

Havell, E.A., Berman, B., Ogburn, C.A., Berg, K., Paucker, K. and Vilček, J.: Two antigenically distinct species of human interferon. Proc. Natl. Acad. Sci. USA 72 (1975) 2185–2190.

Havell, E.A., Yamazaki, S. and Vilček, J.: Altered molecular species of human interferon produced in the presence of inhibitors of glycosylation. J. Biol. Chem. 252 (1977) 4425–4433.

Hayes, T.G., Yip, Y.K. and Vilček, J.: Le interferon production by human fibroblasts. Virology, 98 (1979) 351–363.

Hovanessian, A.G., Brown, R.E., Kerr, I.: Synthesis of low molecular weight inhibitor of protein synthesis with enzyme from interferon-treated cells. Nature 268 (1977) 537–539.

Hovanessian, A.G. and Kerr, I.: The (2'-5')oligoadenylate (pppA2'5'A2'5'A) synthetase and protein kinase(s) from interferon-treated cells. Eur. J. Biochem. 93 (1979) 515–526.

Isaacs, A. and Lindenmann, J.: Virus interference, I. The interferon. Proc. Roy. Soc. B 147 (1957) 258–267.

Johnson, H.M.: Effect of interferon on antibody formation. Texas Repts. Biol. Med. 35 (1977) 359–369.

Knight, Jr., E.: Interferon purification and initial characterization from human diploid cells. Proc. Natl. Acad. Sci. USA 73 (1976) 520–523.

Knight, Jr., E., Hunkapiller, M.W., Korant, B.D., Hardy, R.W.F. and Hood, L.E.: Human fibroblast interferon: Amino acid analysis and amino terminal amino acid sequence. Science 207 (1980) 525–526.

Lebleu, B., Sen, G.C., Shaila, S., Cabrer, B. and Lengyel, P.: Interferon, double-stranded RNA, and protein phosphorylation. Proc. Natl. Acad. Sci. USA 73 (1976) 3107–3111.

Maxam, A.M. and Gilbert, W.: A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA 74 (1977) 560–564.

Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Ecsödi, J., Boll, W., Cantell, K. and Weissmann, C.: Synthesis in *E. coli* of a polypeptide with human leukocyte interferon activity. Nature 284 (1980) 316–320.

Proudfoot, N.J. and Brownlee, G.G.: 3'-Non-coding region sequences in eukaryotic messenger RNA. Nature 263 (1976) 211–214.

Rubinstein, M., Rubinstein, S., Familletti, P.C., Miller, R.S., Waldman, A.A. and Pestka, S.: Human leukocyte interferon: Production, purification to homogeneity, and initial characterization. Proc. Natl. Acad. Sci. USA 76 (1979) 640–644.

Russell, G.J., Walker, P.M.B., Elton, R.A. and Subak-Sharpe, J.H.: Doublet frequency analysis of fractionated vertebrate nuclear DNA. J. Mol. Biol. 108 (1976) 1–23.

Seeburg, P.H., Shine, J., Martial, J.A., Baxter, J.D. and Goodman, H.M.: Nucleotide sequence and amplification in bacteria of structural gene for rat growth hormone. Nature 270 (1977) 486–494.

Schmidt, A., Chernajovsky, Y., Shulman, L., Federman, P., Berissi, H. and Revel, M.: An interferon-induced phosphodiesterase degrading (2'-5') oligoisoadenylate and C-C-A terminus of tRNA. Proc. Natl. Acad. Sci. USA 76 (1979) 4788–4792.

Smith, H.O. and Birnstiel, M.L.: A simple method for DNA restriction site mapping. Nucl. Acids Res. 3 (1976) 2387–2399.

Stewart, II, W.E.: The Interferon System. Springer-Verlag, Vienna, 1979.

Tabak, H.F. and Flavell, R.A.: A method for the recovery of DNA from agarose gels. Nucl. Acids Res. 5 (1978) 2321–2332.

Tan, Y.H., Barakat, F., Berthold, W., Smith-Johannsen, H. and Tan, C.: The isolation and amino acid/sugar composition of human fibroblast interferon. J. Biol. Chem. 254 (1979) 8067–8073.

Taniguchi, T., Sakai, M., Fujii-Kuriyama, Y., Muramatsu, M., Kobayashi, S. and Suda, T.: Construction and identification of a bacterial plasmid containing the human fibroblast interferon gene sequence. Proc. Japan Acad. 55, Ser. B (1979) 464–469.

Taniguchi, T., Fujii-Kuriyama, Y. and Muramatsu, M.: Molecular cloning of human fibroblast interferon cDNA. Proc. Natl. Acad. Sci. USA (1980) in press.

Vilček, J., Havell, E.A. and Yamazaki, S.: Antigenic, physicochemical and biologic characterization of human interferons. Ann. N.Y. Acad. Sci. 284 (1977) 703–716.

Communicated by C. Weissmann.